

PF 18-JUN-1997; 97WO-SF01091.
 XX
 PR 20-JUN-1996; 96SE-0002496.
 XX
 PA (FLOC/) FLOCK J.
 PA (FRYK/) FRYKBERG L.
 PA (GUS/) GUS B.
 PA (LIND/) LINDBERG M.
 PA (NILS/) NILSSON M.
 PI
 XI Flock J, Frykberg L, Gus B, Lindberg M, Nilsson M;
 XX MPI: 1998-063079/06.
 DR N-PSDB: AAV04279.
 XX
 XX Fibrinogen-binding protein from coagulase-negative Staphylococcus
 PT used for prevention, treatment and diagnosis of Staphylococcus
 PT Infection
 XX
 PS Example 3; Fig 6; 45pp; English.

The protein comprises the fibrinogen binding protein (FIG) of coagulase-negative staphylococcus epidermidis HB. Its amino acid sequence was deduced from the isolated fig gene (see AAV04279). The closest known analogue of FIG is the clumping factor of S. aureus which also binds fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG polypeptides can be expressed in host cells. They are used as immunogens, particularly in vaccines (which may be expressed in vivo) to protect humans and animals against coagulase-negative staphylococcus infection. Antibodies raised against FIG can be used for passive immunisation. They block the adherence of bacteria and for diagnosis.

Sequence 1092 AA;

Query Match 100.0%; Score 5646; DB 19; Length 1092;
 Best Local Similarity 100.0%; Pred. No. 1,4e-266;
 Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MINKNNLLTKKKPKANKSNKAIKRFYGTASTIYGATLFLGLGHNEAKAEENSVQDVK 60
 DB 1 MINKNNLLTKKKPKANKSNKAIKRFYGTASTIYGATLFLGLGHNEAKAEENSVQDVK 60
 QY 61 DSNTDELSDSDSDSDDEKNDVNNNOSINTDNNQIIRKKEETNNYDGIKRSDEPTES 120
 DB 61 DSNTDELSDSDSDSDDEKNDVNNNOSINTDNNQIIRKKEETNNYDGIKRSDEPTES 120
 QY 121 TTNDENEAFLQKTPQDNTHLTEEEVKESSESSNSIDTAAQPSHTTINRESVQTS 180
 DB 121 TTNDENEAFLQKTPQDNTHLTEEEVKESSESSNSIDTAAQPSHTTINRESVQTS 180
 QY 181 DNEVESHVSDPANSKIKSNTESGKEENTIDQPNKVKEDSTTSQPSGYTINDEKTSNDE 240
 DB 181 DNEVESHVSDPANSKIKSNTESGKEENTIDQPNKVKEDSTTSQPSGYTINDEKTSNDE 240
 QY 241 LLNLPINEYENKARPLSTSAQPSIKRYTVNQLAEQGSNNVHLIKVDOSTT3CYDSE 300
 DB 241 LLNLPINEYENKARPLSTSAQPSIKRYTVNQLAEQGSNNVHLIKVDOSTT3CYDSE 300
 QY 301 GVKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTPSOLDTSFTIPKIDNGEIIA 360
 DB 301 GVKAHDAENLIYDVTFEVDKVKSGDTMTVDIDKNTPSOLDTSFTIPKIDNGEIIA 360
 QY 361 TGTYNKKKQITTYTVDYDKYENIKAHKLKLTSYIDKSKVPNNNTKLDVEYKTAHSV 420
 DB 361 TGTYNKKKQITTYTVDYDKYENIKAHKLKLTSYIDKSKVPNNNTKLDVEYKTAHSV 420
 QY 421 TTIVEYORPENRTANLQSMFTNIDTKNHTVEQTIYINPLRSARETNVNTSGNDEGST 480
 DB 421 TTIVEYORPENRTANLQSMFTNIDTKNHTVEQTIYINPLRSARETNVNTSGNDEGST 480
 QY 481 IIDDSTIIKVKVGNQNLPSNRIYDYSEEDVTNDDYLAQLGNNDVNINFGNIDSPFI 540
 DB 481 IIDDSTIIKVKVGNQNLPSNRIYDYSEEDVTNDDYLAQLGNNDVNINFGNIDSPFI 540

DB 481 IIDDSTIIKVKVGNQNLPSNRIYDYSEEDVTNDDYLAQLGNNDVNINFGNIDSPFI 540
 QY 541 IKVISKYPNKKDYTTIOQTVMTQTTINEYTGFEFTASDNTIARSTSGGQGGDLPEEK 600
 DB 541 IKVISKYPNKKDYTTIOQTVMTQTTINEYTGFEFTASDNTIARSTSGGQGGDLPEEK 600
 QY 601 TYKIDYVWEDYDKGIONTNDNEKPLSNVLYTLTPPGTSKSVRTDEGKQFGLKNG 660
 DB 601 TYKIDYVWEDYDKGIONTNDNEKPLSNVLYTLTPPGTSKSVRTDEGKQFGLKNG 660
 QY 661 LTKITTFEPEGYPTTLKHSGTNPALDSEGNVWYTINGODMTIDSGFYQIPKISLGNY 720
 DB 661 LTKITTFEPEGYPTTLKHSGTNPALDSEGNVWYTINGODMTIDSGFYQIPKISLGNY 720
 QY 721 VWYDTNKGIOGDEDEKIGSVKVTLKDENGNIISTTTDENCKYQFDMLNSGNYIHFEDK 780
 DB 721 VWYDTNKGIOGDEDEKIGSVKVTLKDENGNIISTTTDENCKYQFDMLNSGNYIHFEDK 780
 QY 781 PSGMTQTITDSDDDDEQADGEVAVTITDHDDEFIDNGYDYDESDSDSDSDSDSDSD 840
 DB 781 PSGMTQTITDSDDDDEQADGEVAVTITDHDDEFIDNGYDYDESDSDSDSDSDSDSD 840
 QY 841 SD 900
 DB 841 SD 900
 QY 901 SD 960
 DB 901 SD 960
 QY 961 SD 1020
 DB 961 SD 1020
 QY 1021 SD 1080
 DB 1021 SD 1080
 QY 1081 LGRKKRNRKNK 1092
 DB 1081 LGRKKRNRKNK 1092

RESULT 2
 ABP40469
 ID ABP40469 standard; Protein: 930 AA.
 XX
 AC ABP40469;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO: 5314.
 DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 XX
 PR 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR MPI: 2002-381255/41.
 DR N-PSDB: ABN93014.

QY 724 DTNKDGIQGDDEKGISGVKVTCLKDENGINISTTTTIDENGKYQFDNLNSGNYIVHFDKPSG 7

XX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
XX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

XX		TM	ECCLES	TT	HOCK	M
PI		Batt+1				

DR MPI: 2000-237781/20.
 DR N-PSDB; AA51202.
 XX Composition used for generating immune response or for inhibiting
 PT microbial colonization in an animal comprises antibodies that bind
 PT collagen binding protein, fibrinogen binding protein and, optionally,
 PT fibronectin binding protein.
 XX Claim 8; Fig 4; 115pp; English.
 XX
 CC The patent discloses multicomponent vaccines containing selected
 CC combinations of bacterial binding proteins termed MSCRAMM (microbial
 CC surface components recognising adhesive matrix molecules) or their
 CC antibodies. A vaccine composition is provided that includes collagen
 CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein
 CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB),
 CC and optionally a fibronectin binding protein e.g. FBP-A.
 CC The vaccines are useful for imparting protection against a broad
 CC spectrum of Staphylococcal strains and for inhibiting microbial
 CC colonisation, especially of Staphylococcus aureus, in an animal.
 CC The combinations can also be used to select donor blood pools for the
 CC preparation of purified blood products for passive immunisation.
 CC The present sequence is a serine-aspartate repeat region
 CC protein, SdrC from Staphylococcus epidermidis. The Sdr protein is
 CC useful in vaccine preparation in combination with specific
 CC bacterial binding proteins. These vaccines can be used to treat a broad
 CC spectrum of bacterial infections, including those arising from both
 CC coagulase-positive and coagulase-negative bacteria.
 CC
 XX Sequence 991 AA:
 SQ
 Query Match 79.4%; Score 4485; DB 21; Length 991;
 Best Local Similarity 81.1%; Pred. No. 3.6e-210;
 Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

Db 574 ISKYPKDDYTTIQGTVMQTINNEYTGFEFRTASDNTIAFSTSSGOGGDLPEPKTYK 633
 QY 604 IGDYWEDEVKDGIGQNTNDEKPLSNVLTLPYDGTSKSVRDEDEKYOFGKNGLT 663
 Db 634 IGDYWEDEVKDGIGQNTNDEKPLSNVLTLPYDGTSKSVRDEDEKYOFGKNGLT 693
 QY 664 KITFETPEGYTPPLKHSCTNPALDSEGNYSWVTINGDDMTIDSGFYQTKYSLGNVWY 723
 Db 694 KITFETPEGYTPPLKHSCTNPALDSEGNYSWVTINGDDMTIDSGFYQTKYSLGNVWY 753
 QY 724 DTNKGIGQDDEKIGSVKYLKDENGNITSTTTDENGKYQPNLNSGNVYHFDKPSG 783
 Db 754 DTNKGIGQDDEKIGSVKYLKDENGNITSTTTDENGKYQPNLNSGNVYHFDKPSG 813
 QY 784 MTQTTDSDGDDDEADAGEEYHVTITDHPDPSIDNGYDESDSDSDSDSDSDSDSDS 843
 Db 814 MTQTTDSDGDDDEADAGEEYHVTITDHPDPSIDNGYDESDSDSDSDSDSDSDSDS 866
 QY 844 DS 903
 Db 867 ----- 866
 QY 904 DS 963
 Db 867 ----- 866
 QY 964 DS 1023
 Db 867 DSDGLD 913
 QY 1024 GSOSD 1083
 Db 914 -----NSSDKTKDKLPDTGANEDHDSKGTLLALFAGLGLLGG 954
 QY 1084 RRRRRKN 1092
 Db 955 RRRRRKN 963

RESULT 5
 AAY08643
 ID AAY08643 standard; protein; 1166 AA.
 XX
 AC AAY08643;
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE S. aureus SdrE protein.
 XX
 KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;
 KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
 KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
 KW extracellular matrix; vascular graft; vascular stent; vaccine;
 KW intravenous catheter; artificial heart valve; cardiac assist device;
 KW antibacterial.
 XX
 OS Staphylococcus aureus.
 XX
 PN W09927109-A2.
 XX
 PD 03-JUN-1999.
 XX
 PE 25-NOV-1998; 98NO-US25246.
 XX
 PR 31-AUG-1998; 98US-0098427.
 PR 26-NOV-1997; 97US-0066815.
 XX
 PA (EIDH/) EIDHIN D N.
 PA (FORE-) FORAS T/A BIORESEARCH IRELAND.
 PA (FOST/) FOSTER T J.
 PA (HOOK/) HOOK M A O.
 PA (INH-) INHIBITEX INC.
 PA (JOSE/) JOSEFSSON E.

PA (PART/) PARTI J M.
 PA (PERK/) PERKINS S E.
 XX
 PI Eldin DN, Foster TJ, Hook MAO, Josefsson E, Partl JM;
 PI Perkins SE;
 XX
 DR WPI; 1999-357844/30.
 DR N-PSDB; AAX77594.
 XX
 PT Staphylococcus aureus fibrinogen-binding proteins for treating
 PT septicemia, osteomyelitis, mastitis or endocarditis
 XX
 XX
 Claim 8; Fig 9; 143pp; English.

This invention describes novel Staphylococcus aureus fibrinogen-binding proteins that bind both the alpha and beta fibrinogen chains. The proteins (and their encoding nucleic acids) are ClfB, SdrC, SdrD and SdrE. Staphylococcus aureus is thought to utilize fibrinogen to adhere to medical devices, binding proteins that bind both the alpha and beta fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as competitive inhibitors to block this binding. Antibodies against ClfB, SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding. The proteins of the invention can be used in a pharmaceutical composition for the treatment of Staphylococcus aureus infection e.g. septicemia, osteomyelitis, mastitis or endocarditis or to inhibit the binding of S. aureus to the extracellular matrix. The proteins or their fragments may be used to coat a medical device to reduce the S. aureus infection of an indwelling medical device, especially where the medical device is selected from the group consisting of vascular grafts, vascular stents, intravenous catheters, artificial heart valves, and cardiac assist devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or encoding gene may be used as a vaccine. The DS (aspartate serine) repeat region or a gene encoding it may be used as an identifying probe for the identification of genes and encoding proteins from Staphylococcus aureus (other than ClfA), S. hemolyticus, S. lugdunensis, and S. schleierferti. The proteins of the invention have antibacterial activity.

Sequence 1166 AA:

Query Match 47.88; Score 2698.5; DB 20; Length 1166;
 Best local similarity 48.2%; Pred. No. 3; 2e-123;
 Matches 595; Conservative 158; Mismatches 267; Indels 215; Gaps 26;

1 MIKKKN-NLTTRKPTANKSNKTAIRKFTYGTASIVGATLLFGLGNEAKAEENSQDV 59
 1 MINKDNKAITKGMISNRLNKFISIKRYTGTASILVGLTFLGLGQEKAAEMT--ST 58
 60 KDSNTDELSDSDSDEEKNVNNOSINTDNNQIKKEETNNYDGIKRESEDRTE 119
 59 ENAKODATSDKKEVSEENNSTENSTNP-----IKKE--TNTDQPEAKKESTS 110
 120 STTNVDENEAFLQKTPQDNTLHLEEVKES---SSVSSNSSIDTAQOPSHTTNEES 176
 111 SSTOKOONNTATTEFPQN--IEKENVKPSTKTATEDSVLLEKKAKNNNT--NNDVT 166
 177 VQTSADVEDHVDFAKSKESNT---ESGKEENTIEQNKYKEDSTTSQPSGYTNIDE 233
 167 TKPS-----TSEPSTSEIQTKPTPOESTNIENSQPOPTPSKVD--NQVDTATPKKE 216
 234 KIS-NODELNLNLP-----INEYENKARPLSTSAQPSIKRV-----TNNOLAEO 277
 217 PVAVSKKEELKNNEPEKLELYRNSDNTDHSKPATAPTAAPKRVAKMFAVAQAPAAVA 276
 277 SNNVNDLITKQTIKYG--GKDNVAALAHGKIDIEYDEFTIINKYKKGDTMINTYDKN 335
 278 GSNVNLIKVTDOSTEGYDSEGVIRAHDAENLIYDVEVDKYSKSGDTMTVDIDKNT 337
 338 VPSDLTDSFTIRIKONSGLIATGTDNKNKQITTYFTYVVKYENIKAHLLSTYDK 397
 336 IPSDLTGDKNPIDITPDSGEVIAKGFDAKATQITTYFTYVVKYEDIKSRLTLYSTDK 395
 398 SKPVNNNTKLDVEYKTALESVNNKTIIVEYORPENRANTQSMETNIDTNNHVEQTIYI 457

396 KTVP-NETSLNLTFAAGKETSONVTVDYQDPMVHGSNQSISFTKLDEDKQTEQDIYV 454
 458 NPLRYSAKKEENNVISGNG-----DESGTIIDSTIIKRYKVGDNQNLPSNRRIYD 508
 455 NPLKRSATNTKVDIASQVDDYDNIKLNGSTIIIDNTEIKRYKVNSDOOLPQSNRIYDF 514
 509 SEYEDVTND-DYAOLGNNNDVNFNGINIDSPYIIKRYKSDPNKDDVTTLQGTVMQTTI 567
 515 SOYEDVTSOPDNKRSFSNNVATLDFGDINSAYIIKRYKSTPPISDGLDIAQGSTMAFTT- 573
 568 NEYTGFEFRAASYDNTIAFTSSGOGGDL--PPEKTYKIGDYVEDVDKDIQNTNDNEKP 626
 574 DKY-GYNNYAGYENFIYTSNDTGGGDTVAPPEKTYKIGDYVEDVDKDIQNTNDNEKP 632
 627 LSNVLTLPDPTDSKSVRRDEDEKQFOFDLKNGLYIKIFEPPEGTPTLKSIGFAPAL 686
 633 MANVLVLTLPDPTDSKSVRRDANGHEFGGLKDEGYTVKFEPTGYLTPKRVNGTIDGK 692
 687 DSGNSVWVTINQDPMPTIDSGFYQFPKYSILGVNWDTKKDGIGQDEKIGSGVETLK 746
 693 DSGNSVTVKNGKDKMSLDTGFYKEPKYMLGDYVEDTKDGIQDANEGIKDYKVTLK 752
 747 DENGNIISTTT----- 757
 753 DSGKVIYGTITTDASGKRYKFTDLNGNVTVEFETPAGYPTVKNTADDKDSNGLFTTG 812
 758 ----- 757
 813 IKDADNMTLDRGFYKTPKYSLAGDYWYDSNKNKDGKODSTEGKINDVYTLQNEKEVIGTT 872
 758 -TDENGKYQPDINSGNVIYHFPKPSGMTOTTTDSGDDKDDADGEEVHTITDHDPSI 816
 873 KTBENGKTYRDNLDGKRYKIFEPKAGLTOTVNTTDD--KDADGCVYDVTITDHDFTL 931
 817 DNGYDESD 876
 932 DNGYFEEDT-----SD 980
 877 SD 936
 981 SD 1040
 937 SD 996
 1041 SD 1100
 997 SD 1055
 1101 SDSDSDGKHTPYKPKMS-----TTKHNNKAKALPE 1131
 1056 TGANEDYGSKGTLLGTLFAGLALL--GKRRKNK 1089
 1132 TGSENNGSNNATLFGGLFALGSLLEFGRKKQWK 1166

RESULT 6
 AAY83170
 ID AAY83170 standard; Protein; 1802 AA.
 XX
 AC AAY83170;
 XX
 DT 24-JUL-2000 (first entry)
 XX
 DE Cell wall protein SdrF.
 XX
 SdrF, SdrG, SdrH: coagulase negative; staphylococcus; septicemia;
 KM osteomyelitis; endocarditis; immune response; vaccine; graft;
 KM stent; intravenous catheter; heart valve; cardiac.
 XX
 OS Staphylococcus sp.
 XX
 FT Key Location/Qualifiers
 MISC-difference 12

[illegible][illegible]

```

XX Sequence 1349 AA:
SQ
Query Match 36.6%; Score 2069; DB 22; Length 1349;
Best Local Similarity 36.6%; Pred. No. 1.6e-92;
Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;

QY 1 MINKKN-NLTKKPKPIANKSKYAIRKTTVGTASTVIGATLFLGIGHNEAKAEKNSVDV 59
DB 1 MLNRENKTAIRKGVNSRLNKFIRKYTGASTVIGTTLIFLGIGNOEAKAEKSTNKL 60
QY KDSNTDELSDSNODSSDEKNDVNNNSINTDNNNOIKKEETNNNDGIEKREDEKTE 119
DB 61 NEATT-----SASDNOSD--KYDMOQLNOEDNTKNDQ-----KEM-----VSSQNETTS 105
QY 120 STTVNDEATFLAKTPODNTHLTEEVKESSESSSIDTAQOPSHHTINRESVQT 179
DB 106 NGNKLIEKES--VQSTGNKKEVEVSTAKSDEQASPSTNEDINTKQ-----TISNCEALQ- 157
QY 180 SDNVEDSHVSDPANKIKESNTESKEENTIEQPNKVKEDSTTSPSGEYTN--DEKISNQ 238
DB 158 PDLOENKSVN-----VOPTNEENKVD-----AKTEST-----LVNKSDAIKSN 198
QY 239 DEL--NLPINEYEKARPLSTTSAP-----SIRKVTNOLAEOGSNVLIKYTDQSI 292
DB 199 DETLVNNSNSNNENNADIILPKSTAPKRLNTRMRIAIVQSPSTAKVNDLITENTLIT 258
QY 293 TEGYDSEGVYKAHDAENLIYDVEFDVKYSGDPTVDDIDKNTV-----PSDLTDSF 346
DB 259 VYDADKNNKIYPAQDYLSLKSQIT--VDDKVSQDYFTIKY--SDTVQYGLNPEIDIKN-- 313
QY 347 TTPKIKD--NSGEIITAGTYDNKKNKOITTFPDYVKNENIKAHKLTSYIDKSKAPNNNT 405
DB 314 -IGDIKDENNGEITATAHDFANMLITYFTDYVDRFNSVQMGINYSIYMDATIP--VS 370
QY 406 KLUDVEKATALSVENTIVEYORPN--ENRATANQSMFNIDTKNHT-----VEOTI 455
DB 371 KNDVEFNNTIGTTTKTANIOYPRYVNEKNSIGSAFT--ETVSHVGNKENKPGYKOTI 428
QY 456 YINPLRYSAKETNNVI-----SCNGDEGSTITIDSTIIKVKYKGNQNLPSNFIYDYS-- 509
DB 429 YNPPSENSLTNKLKLVQVHSSYPNNIGINKDVTDIKIYQPKGYTL--NKGDVNTK 485
QY 510 EYEDVTNDYIAO--LGNNDVNIINFGNIDSPYIIKIVISKYPNKDDITTOQTYTMOIT 566
DB 486 ELTDTN--OYLQKITTYGDNNSAVIDFGNADSAVYVWMTKRFQYTNSESEPTLVQMATLST 544
QY 567 INEYGEFRTASVDNTIAFSTSSGQGDLPPEKTYKIDYVWEDVDKDGIO----- 618
DB 545 GN-----KSVSTGNALGFTNNQSGAG-----QEVYKIGNYVWEDTKNKGVELEGEKGVG 594
QY 619 -----NTN----- 621
DB 595 NVTVTFPDNNTNTKYGEAVTKEDGSYLLPFLPNDGYRVEFSLPKGYEVTPSKQNNBEL 654
QY 622 -----DN-----EKLPSNVLVTL-- 634
DB 655 DSNGLSSVTVNGKDNLSADLGIYKPKYNLCDYVWEDTNKNGIODODEKIGISGVTVTLKD 714
QY 635 -----TYPDC----- 639
DB 715 ENGVDLVKVTYTTADGKRYKTTDLHNGNKYKEFTTPEGYTPTTYTSSGSDIEKDSNGLTTTGV 774
QY 640 -----TSKSV----- 644
DB 775 INGADNMTLDSGFYKPKYNLGNVYWEDTNKDKGKODSTEGKISGVTVTLKNNGEVLQTT 834
QY 645 RPDDEGKYOFDLKNGLTIKITFERPEGTYTFLKHSGTNPALDSEGS----- 692
DB 835 KTDKDGKYOTGLENG--TYKVEFEPSGYTPVOVSGTDEGIDSNCTSTTGVIKDKDNDT 893
QY 693 -----VW-----VTI-----NGQ-- 700

```

```

DB 894 IDSGFYKPTYNLGDYVWEDTNKNGVQDKDEKIGISGVTVTLKDENDKVLKTYTTDENGKYQ 953
QY 701 -----DMTIDSGFYOTPK 714
DB 954 FTDLNGTYKVEFEPTSGYTPTSVSGNDTEKDSNGLTTTGVIKADNMTLDSGFYKPK 1013
QY 715 YSLGNVYVDTNNKQIQGDDDEKIGISGVTVTLKDENGNIISTTTDENGKYQFDNLSNGY 774
DB 1014 YSLGVYVYDNNKQDKODSTEGIKNDVKTTLNKEGVIGTTKTDENKRYCFDNLDSGY 1073
QY 775 IVHEDKPSGMYQTQTTSDGDDDEQADAGEEVHTYTHDDFSIDNGYVDESPSDSDSDSD 834
DB 1074 KVIFERKPAGLQTYNTTEDD--KQADGGEVYTYTHDHDFTLNDKYPEDT----- 1123
QY 835 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 894
DB 1124 --SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1181
QY 895 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 954
DB 1182 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1241
QY 955 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1014
DB 1242 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1300
QY 1015 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1073
DB 1301 -----TTKDHNNKAKALPETGSENNGSNNATLFGLE 1332
QY 1074 AGLGALL-GKRRKNRK 1089
DB 1333 AALGSLLEGRKKONK 1349
RESULT 9
AA037544
ID AA037544 standard; Protein: 1349 AA.
AA037544:
14-FEB-2002 (first entry)
Staphylococcus aureus cellular proliferation protein #1714.
DE Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
OS Staphylococcus aureus.
XX WO200170955-A2.
FN
XX 27-SEP-2001.
PD
XX 21-MAR-2001; 2001WO-US09180.
PF
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
PA
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS55403.
XX
XX New polynucleotides for the identification and development of

```

PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Example 3; Seq ID No 13137; 511bp; English.

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPRO at
 CC ftp.wipro.int/pub/published_pct_sequences.

CC Sequence 1349 AA;

Query Match 36.6%; Score 2069; DB 22; Length 1349;
 Best Local Similarity 36.6%; Pred. No. 1.6e-92;
 Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;

QY 1 MIRKN-NLTKKKPIANKSNKAIKRTVGTASIVGATLLFGHNEKAKEENSQDV 59
 DB 1 MLNREKTAITRKGMSNRNKSIRKTYGTASILVGLTILFGLNGEKAESTKEL 60
 QY 60 KOSNTDELSDSQSDSEKNDYNNOSINTDNNQIITKEETNNYDGIKESGERT 119
 DB 61 NEATY-SASDNOSD-KVDMQOQEDNTKNDNQ-KEM-VSSQGNETTS 105
 QY 120 STNVENETFELOKTPQDTHLTHEEVKSSSVESNSSIDPAQPSHTTINEEVSOT 179
 DB 106 NGKKLLEKES-VOSTGKNVEVSTAKSDQASPKSTNEDLNTKO-TISQELAQ- 157
 QY 180 SDVDESHVSDPANSKIKESSTESGKEENTIEQPNKYEDSTTSQPSGYTNI-DEKISNQ 238
 DB 158 PDLQEKSVNV-VQPTNEENKAYD-AKTESTT-LNVKSDAKSN 198
 QY 239 DELL-NLPINEXENKARPLSTSAOP-SIKRYTVNOLAEOGSNVHLKVTQDSI 292
 DB 199 DELTVNNSNSNNENNADILPKSTAPKRLNTRMRIRAAVOSSPEAKNVNDLITSNTLT 258
 QY 293 TEGYDSEGVYKAHDAENLYDYTFEYDDVKYSGDTMTVIDKNTV-PDLTDSF 346
 DB 259 VVDADKNKTVPAQDYLSTKSOIT-VDDKVKSGDYFTIY-SOTVGVGLNPDINK- 313
 QY 347 TIRKID-NGSEIATGTYNKKNQIYTFYTDYDKENIKAHKLISYIDSKVPMNNT 405
 DB 314 -IGDIDPNNGEITATRAKHTANLLITFTDYDRNSQMGNTSYIYDADIP-VS 370
 QY 406 KLDVEKTAISVNVKTIYEQRPN-ENRPAIQQSMFTNIDTKNHT-VEQTI 455
 DB 371 KNDVEFVNTIGNTTATTANIQPDYVYVNEKNSIGSAFT-ETVSHVGNKENPQYQTI 428
 QY 456 YINPLRYSAKETVNI-SGNGDESGTIIIDSTIKYKVGQDNQLPDSNRIDYS- 509
 DB 429 YVNPSENSLTNAKLVQAVYSSYPNNIGQINKDVTIKITIQVPGYTL-INKGIDVNTK 485
 QY 510 EYEDVYNDYAO-LGNNDVNIIFGNIDSPYIIRKYSKYDPKNDYTTITQVTVQQT 566
 DB 486 ELDDVYN-QYLQKITGDNNSAVIDFENADSAYVWNTFQYNTSSPILYQATIST 544
 QY 567 INEYTGFRASDNTASTSSGOGGLDPEKTYRIGDYWEDVDKQIC-618
 DB 545 GN-KSVSTGNALGFTNNOSGAG-DEVYKIGYVWEDTKNGVQELGKGVG 594

QY 619 -----NTN----- 621
 DB 595 NNTVTFEDNNTNTRKVGAVTKEDGSYILPMLPGNDYVRESNLPKGEYVPSKGNNEEL 654
 QY 622 -----DN----- 634
 DB 655 DSNGLSSVTNKGKDNLSADLGIYKPKYCNLDYWEDTNKNGIODEKIGISGVTVTLKO 714
 QY 635 -----TTPDG----- 639
 DB 715 ENGDIKTVTTDADGKXKFTDLHNGNYKVEFTTPEGTPPTVSGSDIEKDSNGLATTGV 774
 QY 640 -----TSKSV----- 644
 DB 775 INGADNNTLDSGFYKTPKYNLGNVWEDTKKDGKODSTEGISGVYVTLKNGEVLQTT 834
 QY 645 RTDEDKRYOPDGLKNGLYKITEETPEGYPTLKHSCTNPALDSEGS----- 692
 DB 835 KTDKDGKYGFTGLENG-TYKVEFETPGYPTQVSGCTDEGIDSGNSTGTGVIKDKNDOT 893
 QY 693 -----VA-----VTI-----NQ- 700
 DB 894 IDSGFYKPTYNLGDYWEDTNKNGVODKDEKIGISGVTVTLKENDKYLKTVTTDENGKQ 953
 QY 701 -----DDMTIDSGFYQTPK 714
 DB 954 FTDLNNGTYVEFEPTPGYPTSVTSGNDTEKDSNGITTTGVIKADNMTLDSGFYKTPK 1013
 QY 715 YSLGNVWYDPTNKDGIQDDEKIGISGVTVTLKENDNGNISTTTDENGKYOQFNLSGNY 774
 DB 1014 YSLGDYVWYDSNNKDGKQDSTEGIKQDYKVTLLNEKGEVIGTTKIDENGKCYCFOLDGKY 1073
 QY 775 IYHFDKSGMTQTTTDSGDDDEQDADGEVYVYTTTDDPSIDNGYVDDSDSDSDSDSD 834
 DB 1074 KYIFKEPKAGLTQYVYNTTEDD-KDADGEVYVYTTDHDFTLDNGYEEDT----- 1123
 QY 835 SDS 894
 DB 1124 --SD 1181
 QY 895 SDS 954
 DB 1182 SDS 1241
 QY 955 SDSVSD 1014
 DB 1242 SDS 1300
 QY 1015 SDS 1073
 DB 1301 -----TTKDHNNKAKALPETGSENNGSNNATLTfGGLF 1332
 QY 1074 AGLGALL-GRRKRRK 1089
 DB 1333 AALGSLLFGRKKQNK 1349

RESULT 10
 AAU34082
 ID AAU34082 standard; Protein; 932 AA.
 AC AAU34082;
 DT 14-FEB-2002 (first entry)
 XX
 XX
 XX
 DE Staphylococcus aureus cellular proliferation protein #358.
 DE Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 OS Staphylococcus aureus.
 XX


```

PN WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207272P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
XX Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
XX N-PSDB; AAS51941.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 5578; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 932 AA:
XX
XX Query Match 35.6%; Score 2011; DB 22; Length 932;
XX Best Local Similarity 42.9%; Pred. No. 6.6e-90;
XX Matches 481; Conservative 141; Mismatches 278; Indels 220; Gaps 32;
XX
XX 2 INKNNLTKKKRIANKSNKYAIRKFTVGTASTVIGTTLFGGLHNKAKAEISYDVKD 61
XX 1 MNKKATATNKKRGIPIPNLNKFSIRKYSVGTASILVGTLLFGGLSGHAKAAE----- 52
XX
XX 62 SNTDELSDSDSDSDDEKNDVINNNOSINTDNNQIIRKEETNNYGIKRSSEDTTEST 121
XX 53 -HTNGEL-----NOSKN-----ET 65
XX
XX 122 TNDVNEATFLQTPDNTLHTEEVKSSSVSSSIDTAQOPSHTTINRESVQTS 181
XX 66 TAPSEKNT-----EKVDSRQ-----N 83
XX
XX 182 NVEDSHVSDPANKIKESNTESGKEENTIEDPNKVKEDSTTSOPSGTINDEKISNDEL 241
XX 84 NIDOSTSD--QPKVNESDSTVKE--TTEP---ONTSTOPTKONN--DAMAND-- 131
XX
XX 242 LNPINYEKAKRPLSTSNAPSIK-----RVYVNOAA--EGGSNVHLKVTQDOST--- 292
XX 132 -NLAQONISTQADKDVSTTPPTTIKPTLNRMNAVNTVAAPQDGTINADKVFHTNIDIAID 190

```

```

QY 293 -----TEGYDSEGVYKAHDAENLIYDTFEVDKVKSGDTMTVDIDKNTVPSDL 342
XX 191 KGHVNTKTGTETFEWATSDVLK-----LKANYITIDSVKSGDITFEYGGYPRPGSV 242
XX
XX 343 TDSFTIPKIKDNGSEIITAGTYDNKKQIITYFTDYVYKYEINIAHLKILSYIDKSKVFN 402
XX 243 RLPSQTONLYNAGNIIAKIYDSESTTYTFTNVYDQGTINISGSEPEYAFAPKREWAT 302
XX
XX 403 NNTKLDEYKTAASSVNTKITTEYORPENRNTANLQSMFINDTKNHTVQTIYIN-PLR 461
XX 303 DKTAPEMEVTLGNDKYSKNVIYD---GNOKGOQLISTSNVINNEDLSRMTYVYVQPK 359
XX
XX 462 YSAKETNV-NISGNGDEGSTIIDSTIIKVKYKGDNON-----LPDSNRIYDSEEDV 514
XX 360 TYTKEFTFVNLT-----GYFNPDAKNFKIYEV-TNQNQVDSFTPDTSKLTIVTOKFKI 413
XX
XX 515 TNDVYQOLGNNDVYNIFG--NIDSPYIIKVIKYPNKKDYTTIOQTYTMOFTTNEYTG 572
XX 414 T--YSDNKTATVADLNLGSSSDKOYIIIOQVAVPNNSTDNCKIDYLTLETQNGKSSMSN 470
XX
XX 573 EFRFASDNTIATSTSSGQGGDLPEPKTYKIGDYVWEDVDKXGIONTNDNEKPLSNVLY 632
XX 471 -----SYSNVNGSSIANGD-----OKYINLSDYVWEDTNKDKODA--NEKGIKGYV 516
XX
XX 633 TLTYPDGTS-KSVRTDEGKYQFPGDKLNGLYTKITRETPEGYPTPLKHSCTNPALDSEGN 691
XX 517 ILKDSNGKELDRFTTDENGKYQFTGLNG--TYSEFSTLAGYPTTANAGTDAVDSGL 575
XX
XX 692 SWATYTINGDDMTIDSGFYTPKYSIGNYVYDTNKKDGIQDDEKIGSKYKVLKDENGN 751
XX 576 TTGVKIKADANMLDGSFYTPKYSIGDYVWYVSNKNGKDSSTEGKIDYKVLNKEGE 635
XX
XX 752 IISTTTDENGKYQFPGDKLNGLYTKITRETPEGYPTPLKHSCTNPALDSEGN 811
XX 636 VIGTITKDENGKTRFNDLDSGKRYKVFEXFAGLTQISTNTTEDD-KRADGGEVDYITTDH 694
XX
XX 812 DDFSIDNGYVDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 871
XX 695 DDTTIDNGYVEEFT-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 741
XX
XX 872 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 931
XX 742 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 801
XX
XX 932 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 991
XX 802 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861
XX
XX 992 DSDSDSDSDSDSDSDSDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1051
XX 862 DSDSDSDSDSDSDVGHKHFAPKPM-----TYVDHNNKA 892
XX
XX 1052 K-LPDTGANEDYSGKLLGTTFAGIGALLL-GKRRNRK 1089
XX 893 KALPETGSENNNSNNGTLFGGLPALGSLLLFGKRRKQNK 932
XX
XX RESULT 11
XX ID AU036845 standard; Protein; 932 AA.
XX AC AU036845;
XX DT 14-FEB-2002 (first entry)
XX DE Staphylococcus aureus cellular proliferation protein #1015.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.

```


OS Staphylococcus aureus.
XX
PN MO927109-A2.
PD 03-JUN-1999.
XX
PF 25-NOV-1998; 98MO-US25246.
XX
PR 31-AUG-1998; 98US-0098427.
XX
PR 26-NOV-1997; 97US-0066815.
XX
PA (EIDH/) EIDHIN D N.
PA (FOR-) FOREAS T/A BIORESEARCH IRELAND.
PA (FOST/) FOSTER T J.
PA (HOOK/) HOOK M A O.
PA (INH-) INHIBITEX INC.
PA (JOSE/) JOSEFSSON E.
PA (PATT/) PATTI J M.
PA (PERK/) PERKINS S E.
XX
T EIdhIn DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
Perkins SE;
..4
DR WPI: 1999-357844/30.
DR N-PSDB: AAX7592.
XX
PT Staphylococcus aureus fibrinogen-binding proteins for treating
PT septicemia, osteomyelitis, mastitis or endocarditis
XX
PS Claim 8; Fig 7; 143pp: English.
XX
CC This invention describes novel Staphylococcus aureus fibrinogen-binding
CC proteins that bind both the alpha and beta fibrinogen chains. The
CC proteins (and their encoding nucleic acids are Clfb, Sdrc, Sdrc and
CC Sdrc). Staphylococcus aureus is thought to utilize fibrinogen to adhere
CC to medical devices, binding proteins that bind both the alpha and beta
CC fibrinogen chains (Clfb, Sdrc, Sdrc and Sdrc) can therefore be used as
CC competitive inhibitors to block this binding. Antibodies against Clfb,
CC Sdrc, Sdrc and Sdrc inhibit Clfb, Sdrc, Sdrc and Sdrc mediated binding.
CC The proteins of the invention can be used in a pharmaceutical composition
CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
CC S. aureus to the extracellular matrix. The proteins or their fragments
CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. Clfb, Sdrc, Sdrc, Sdrc, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate setine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than Clfb), S. hemolyticus, S. lugdunensis, and S. schlerferi.
CC The proteins of the invention have antibacterial activity.
XX
SO Sequence 930 AA.
Query Match 35.2%; Score 1990; DB 20; Length 930;
Best Local Similarity 42.3%; Pred. No. 6.9e-89;
Matches 462; Conservative 154; Mismatches 293; Indels 184; Gaps 27;
QY 2 INKNNLLKKRPDIANKSKYARKRTVTASTVIGATILFLGIGHAEAKAEKSVQDWD 61
DB 1 MNKKKTATNRKGIPIRLKRFISIRKYSVGTASTLGTLLFGLSGHEAKAAE----- 52
QY 62 SNTDELSDNSDSDSDDEKNDVNNOSINTDNNNOIIEKEETNNVNGEKEKREEDTEST 121
DB 53 -HTNGELNSKNETTAPSENKTKTKKVDKQRLKDNTO-----T 88
QY 122 TYNDENEATFLQTPDNTHTLEEVKESSVESSNSIDTAQPSHTTINRESVQTS 181
DB 89 ATADDPKVT-----MSDATVKESSNM---QSPQNAVANQSTKTSNV 129
QY 182 NVEDSHVDFANSKIKESNTESGKEENTIEQPKVKVEDTSTQPSGYTINDEKISNDEL 241

DB 130 TYNDEKSTTYSNETDKSNLQA-----KDVSTT----- 157
QY 242 LNLPINEYENKARPLSTSAQPSIKRVTYNQOLA-EDGSVNNLHKYV--DOSITEGY-D 297
DB 158 -----PKTTTIPRLNRNAVNTVAPOQGTWNKVKHESNIDIAIDGHVN 204
QY 298 DSEGVK--AHDAENLIYDVEFDKVKSGDTMYDIDKNTVPSDLTDSFTIPKIDNS 355
DB 205 QGTGKTEFNATSSDVLKLNKANYIIDSVKEDGFTTFEYGGYFPRGSVRLSOTQNLNAQ 264
QY 356 GELIATGTDNKKQOTTTFTDYVDKYEIKAKLKLSTYDOKSVNNNTKLDVEKTLAL 415
DB 265 GNIIANGIDSTNTTFTYFTNVQYTVNVRGSEQVAFKRNATTDKATVMEVTLGN 324
QY 416 SSVNTITVEYORPENRNPALQSMFTNIDTKHATYEQTYIN-PLRSKAKETNV-NISG 473
DB 325 DTYSSEIYD---GKKKAPLSSSTNYINNEDELSRMTATVYVQPKNTTYKQTFVNLNLT- 380
QY 474 NGDEGSTIIDSTTIIRYKVDNQNLT---PSNRITYDYSEYEDVYNDYAOLGNNNDV 528
DB 381 ---GYKFNPNAKFKIYEVTQNOQFVDSFTPTSKLKVDYDQFDVY---YSNDNKTATV 453
QY 529 NINFGNIDS--PYIIVY-----SKYDPNKDQYTTIQOYVTTMOTITINETGERTASYDN 561
DB 434 DLKKGOTSSNKQYIIQOVAVPDNSSTDNCKIDY-----TLDTDKTKYSW---SNSYSN 483
QY 582 TIAFTSSGCGQGDLPPEKTYIGDYVMDVDKDIQNTDNKPKLSNVYVTLTPPGTS 641
DB 484 VNSSSTANGD-----QKYNLGDVYEDTNKDGKODA--NEKGIKGYVIIKDSNGKE 534
QY 642 -KSVRTDEDKQYFDGLKNGLIYKIFETPEGYPTLKHSTNPALDSEGSVWYVINGQ 700
DB 535 LDRTTDENGKQYFTGLSNG-TYSVESTPAGYTPPTANGTDAVDSOGLTTGVYKDA 593
QY 701 DMTITSGYQTPKYLYGANYWDTNKGDIQGDDEKIGSVKVTYLDENGNITSTTDE 760
DB 594 DNMTTLSGFKTPKYSLGDVWYDSNKGKRDSTENGKIKVYVTLQNEKEVEVIGTETDE 653
QY 761 NGKYOPDNLSNGYIVHFDKPSGMQTTPDSGDDDDADAGEVHYHTPHDFFSINDCY 820
DB 654 NGKRYRDNLSGKYKIFEPKAGLTGTGNTTDD-KDADGGEVDVYTDHDPFTLLNGY 712
QY 821 YDDESDDSD 880
DB 713 YEBET-----SD 765
QY 881 SD 940
DB 766 SD 825
QY 941 SD 1000
DB 826 SD 885
QY 1001 SD 1060
DB 886 SDAGKMTPAKPSV-----KDGKHTKAKALPRTGLEN 917
QY 1061 DYSGKTLTGTF 1073
DB 918 NNSNNGTLFGGLF 930
RESULT 13
AA08642
ID AA08642 standard; Protein; 1315 AA.
XX
AC AA08642;
XX
DT 09-AUG-1999 (first entry)
XX
DE S. aureus Sdrc protein.

CC programmes. The antisense nucleic acid sequence is also useful to screen

[illegible]

698 DSPDSPDSPDSPDSPDSPDSPDSPDSPDSPDSPDSPDSPDSPDSPDSPSPDS 75

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01; Search time 17.6129 Seconds

(without alignments)
1824.221 Million cell updates/sec

Title: US-09-147-405b-15

Perfect score: 5646

Sequence: 1 MINKNNLLTKRRKPIANKSN.....FAGLCALLGKRRKRRKKN 1092

Scoring table:

BLOSUM62

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCROS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4490	79.5	930	4	US-09-134-001C-5314
2	1628	28.8	933	3	US-08-293-728-2
3	1628	28.8	933	4	US-09-421-868-2
4	635.5	11.3	2137	4	US-09-134-001C-4463
5	482	8.5	466	4	US-09-134-001C-4749
6	377.5	6.7	1183	4	US-09-134-001C-3530
7	349	6.2	345	4	US-08-856-253-7
8	331.5	5.9	1161	4	US-09-327-536-2
9	315.5	5.3	1112	2	US-08-714-402-2
10	301.5	5.3	1231	3	US-08-904-263A-4
11	286	5.1	1060	4	US-08-911-393-2
12	282	5.0	3052	2	US-08-557-122A-26
13	282	5.0	3052	4	US-09-262-666-26
14	279	4.9	3696	4	US-09-134-001C-5080
15	278	4.9	699	4	US-09-134-001C-4054
16	276.5	4.9	1085	1	US-08-431-080-28
17	276.5	4.9	1085	2	US-08-938-534-28
18	276.5	4.9	1085	4	US-09-345-294-28
19	276	4.9	2391	2	US-08-446-855A-2
20	276	4.9	2391	4	US-09-150-741-2
21	271	4.8	1115	2	US-08-568-459A-2
22	271	4.8	1115	2	US-08-487-826B-2
23	271	4.8	1115	2	US-09-210-288-2
24	271	4.8	1115	6	5198347-6
25	270.5	4.8	1833	4	US-08-621-944A-4
26	270.5	4.8	1833	4	US-08-945-567D-4
27	270.5	4.8	1992	4	US-08-621-944A-3

28	270.5	4.8	1992	4	US-08-945-567D-3	Sequence 3, Appl1
29	269.5	4.8	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
30	267	4.7	2123	4	US-08-968-685A-10	Sequence 10, Appl
31	265	4.7	2048	4	US-09-268-347-49	Sequence 48, Appl
32	263.5	4.7	2314	4	US-09-268-347-49	Sequence 49, Appl
33	260	4.6	1183	2	US-08-447-031A-2	Sequence 2, Appl1
34	259.5	4.6	1435	2	US-08-568-459A-4	Sequence 4, Appl1
35	259.5	4.6	1435	2	US-08-487-826B-4	Sequence 4, Appl1
36	259.5	4.6	1435	4	US-09-210-288-4	Sequence 179, App
37	257	4.6	1776	4	US-09-556-877-179	Sequence 179, App
38	257	4.6	1776	4	US-08-973-462-8	Sequence 8, Appl1
39	252.5	4.5	1786	4	US-07-667-276A-4	Sequence 4, Appl1
40	252	4.5	821	1	US-09-556-877-195	Sequence 195, App
41	250.5	4.4	821	4	US-09-620-412C-195	Sequence 195, App
42	250.5	4.4	821	4	US-09-620-412C-195	Sequence 4568, Ap
43	250	4.4	690	4	US-09-134-001C-4568	Sequence 70, Appl
44	249.5	4.4	905	4	US-09-074-658-70	Sequence 5, Appl1
45	245	4.3	1850	4	US-09-620-093A-5	

ALIGNMENTS

RESULT 1
US-09-134-001C-5314
Sequence 5314, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5314
LENGTH: 930
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match	79.5%	Score 4490;	DB 4;	Length 930;
Best Local Similarity	81.2%	Pred. No. 1.8e-241;		
Matches 884;	Conservative 19;	Mismatches 26;	Indels 160;	Gaps 3;
QY	4	KNNLLTKRRKPIANKSNKVAIRKFTVGTASIVIGATLLFGLGHNKAKEENSVOYDKSN 63		
DB	2	KNNLLTKRRKPIANKSNKVAIRKFTVGTASIVIGATLLFGLGHNKAKEENSVOYDKSN 61		
QY	64	TDELDSDNSQSDDEKNDVYNNNSQNTDNNQITKEETNYVDGIEKESDEPTESTN 123		
DB	62	MDDELDSDNSQSDDEKNDVYNNNSQNTDNNQITKEETNYVDGIEKESDEPTESTN 120		
QY	124	VDNEATFLOKTDQDNTHLFEFEVKESSEYSSSIDFQOQPSHTTINEESVOYDPN 183		
DB	121	VDNEATFLOKTDQDNTHLFEFEVKESSEYSSSIDFQOQPSHTTINEESVOYDPN 180		
QY	184	EDSHVDSFANSKIKESNTESGKEENTIEOPNKYEDSTSPSGYTNIDEXISNOBELN 243		
DB	181	ENSRVDSFANSKIKESNTESGKEENTIEOPNKYEDSTSPSGYTNIDEXISNOBELN 240		
QY	244	LPINEYENKARPLSTSAOSIKRVYNQAAOGSVNHLIVTDOSTIEGVDSDSGVI 303		
DB	241	LPINEYENKARPLSTSAOSIKRVYNQAAOGSVNHLIVTDOSTIEGVDSDSGVI 300		
QY	304	KADAENLIVDYFEVDKDKKSGDTMTVDIDKNTVPISDLTDSFTIPKIKNSGEIATGT 363		
DB	301	KADAENLIVDYFEVDKDKKSGDTMTVINDKNTVPISDLTDSFTIPKIKNSGEIATGT 360		

QY	364	YDNNKQITFTTFDNDYDKXENIKAHKLTLSYIDSKSVPPNNKTKLDEVEKTKTLSSVNNKIT	423
Db	361	YDNNKQITFTTFDNDYDKXENIKAHKLTLSYIDSKSVPPNNKTKLDEVEKTKTLSSVNNKIT	420
QY	424	VEYQRPENRNTALQSMFTNIDFKNNHTVEQTYIINPLRISAKETNVNISGNGDEGSTIID	483
Db	421	VEYQRPENRNTALQSMFTNIDFKNNHTVEQTYIINPLRISAKETNVNISGNGDEGSTIID	480
QY	484	DSIIIKVYKVGDNQNLPSNRIYDSEYEDVYNDVYAOLGNNDVNINFGNIDSPYIIKV	543
Db	481	DSIIIKVYKVGDNQNLPSNRIYDSEYEDVYNDVYAOLGNNDVNINFGNIDSPYIIKV	540
QY	544	ISKDKPKKDYTTIIQOTVYMQTTINNEYTEEFRTASVDNTIAFSTSGGCGGDLPEPEYTK	603
Db	541	ISKDKPKKDYTTIIQOTVYMQTTINNEYTEEFRTASVDNTIAFSTSGGCGGDLPEPEYTK	600
QY	604	IGDVGWEDVVKDGIQNTNDNKEPLSNVLTLYIPQSTSVKSVPTDSDGKYOFDTGLKNGLT	663
Db	601	IGDVGWEDVVKDGIQNTNDNKEPLSNVLTLYIPQSTSVKSVPTDSDGKYOFDTGLKNGLT	660
QY	664	KITFETPEGYPTPLKHSFTNPALDSGNSVWYTINGQDMTIDSGFYOTPKYSIGNVYVY	723
Db	661	KITFETPEGYPTPLKHSFTNPALDSGNSVWYTINGQDMTIDSGFYOTPKYSIGNVYVY	720
QY	724	DTNKDGIQGDDEKIGISGVKVTYTKDENGNIISTTTDENGKYOFDNLNSGNTIYVHFDKPSG	783
Db	721	DTNKDGIQGDDEKIGISGVKVTYTKDENGNIISTTTDENGKYOFDNLNSGNTIYVHFDKPSG	780
QY	784	MTQTTTSGDDDEQDADAGEVYHTIIPDHQFSIDNGYXDSESDSPSDSDSDSDSDSDSDS	843
Db	781	MTQTTTSGDDDEQDADAGEVYHTIIPDHQFSIDNGYXDSESDSPSDSDSDSDSDSDSDS	833
QY	844	DSDSDDS	903
Db	834	DSDSDDS	833
QY	904	DSDSDDS	963
Db	834	DSDSDDS	833
QY	964	DSDSDDS	1023
Db	834	DSDSDDS	880
QY	1024	GSDSDSDSDSDSDSDNDNLGNSSDKSTDKLPTGANEEDYGSKGTLLGTFLAFLJALLLGK	1083
Db	881	GSDSDSDSDSDSDSDNDNLGNSSDKSTDKLPTGANEEDYGSKGTLLGTFLAFLJALLLGK	921
QY	1084	RRKNRKNKN 1092	
Db	922	RRKNRKNKN 930	
<p>RESULT 2</p> <p>US-08-293-728-2</p> <p>Sequence 2, Application US/08293728D</p> <p>Patent No. 6008341</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Foster, Timothy J.</p> <p>APPLICANT: McDevitt, Damien L.</p> <p>TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene</p> <p>FILE REFERENCE: 05344.105011</p> <p>CURRENT APPLICATION NUMBER: US/08/293, 728D</p> <p>CURRENT FILING DATE: 1994-08-22</p> <p>NUMBER OF SEQ ID NOS: 20</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 2</p> <p>LENGTH: 933</p> <p>TYPE: PRT</p> <p>ORGANISM: Staphylococcus aureus</p> <p>US-08-293-728-2</p>			

Best Local Similarity	37.6%	Pred. No. 9.5e-83			
Matches 420:	Conservative 153:	Mismatches 313:	Indels 230:	Gaps	35:
QY	18 KSNKAIKRFYTGTSIYIGATLLG-IGHNEKAENSVDQKDSNTDDELSHDQSS	76	1	1	1
Db	5 KKEKAIKRSKISGVASVYLGTIGGLSSKRADASSENSV-----TQSMASAN	52	1	1	1
QY	77 DEEKDVINNNOSINTDNNQIILKKEETNNYDGIKRESDRETESTTNDENATEFLQKTP	136	1	1	1
Db	53 ESKSDSSSVSAAPPTDDTNV-----SDTKSSNTN---NETSYAQP	93	1	1	1
QY	137 QDNHILTEEEVKESSSVSSNSIDTAQQPSFTTINRESVQTSIDNVESHVSDPANSKI	196	1	1	1
Db	94 -----AQOETTOSSSTNATTEETPVLTGATTTTNNQANTPATOS-SMTNAEELVN---	143	1	1	1
QY	197 KESNTESGKEENTYIOPKNVKEDSTTSQSGYTIINDEXISNODELLN-LPINEYEKARP	255	1	1	1
Db	144 QTSNETTENDTNV-----SSVNSQNSNTMAENSTTODDTSEATPSN---NESAP	191	1	1	1
QY	256 LSTTSAPSIKRYVN-----QLAAEGSNVNHLLIKYTDQ-SITEGYDSEGVIK	304	1	1	1
Db	192 QSTDSANNDVYNAQVNTSAPRRARAESLAVAADAPAGADTINQLTNLYVGI-DSGTYIV	250	1	1	1
QY	305 AHDAENLIYDYFEYVDKKSGDPTMYDIDKNTVPSDLSTPIRIKNSGEIILATGY	364	1	1	1
Db	251 PQAQGVKLNLYGFSPVNSAVKGDTRKITVPKELNLNGVSTAKVPIIMAGD-QVLANGVI	309	1	1	1
QY	365 DRKNQIITTFDYDYDKYENIKAHKLSTGISDKSKVPNNNTKL-DEYKTALSS--VNYT	421	1	1	1
Db	310 DS-DGVNITTFIDYNTKDKVAKATLIMPAYID---PEVVKTKGNTLATLGIGSTANT	364	1	1	1
QY	422 ITVEYQRPENRNTANLQSKFTYIDTKNHTVEQTIYNPL-RYSAKETVNVISGNDGSS	479	1	1	1
Db	365 VLADYEKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDVNIAPVLGNLKPNTDSNA	424	1	1	1
QY	480 TIIIDSTIIKKYKVGDNQMLPDSNRIYDSEXEDVNDNDYQAQGNNDVNINGC---NI	535	1	1	1
Db	425 LIDQQTSTIKYKVDAAALSESIFY-NEPNEEDVYNSVNIITPPNPQYKVEFNTDDQI	483	1	1	1
QY	536 DSRPIIKYISKDPN-KDQYTTIQQTVYMQTINNEYTGE-FRTASYDNTIARSTSGGQ	592	1	1	1
Db	484 TPIYIVVNGHIDPNKGD-----LALRSTLYGVNSNIMISMWMDNEVAANNNGSGS	536	1	1	1
QY	593 QGDLPPEKTYKIGDYWEDVDKDGIONTDNDEKPELSNVLYTLTPDGTSKSVRTDEGKY	652	1	1	1
Db	537 DG-----IDKPYVP---EQP-----DEPEI	554	1	1	1
QY	653 QTDGLKNGLYTKITFEETPGCYPTPLKHSSTNPALDSEGSVWYTINGQDDMTIDSGFYQT	712	1	1	1
Db	555 E-----PIPEDSSDDPSGSDSGSS-----NSDGSGSDSGST	586	1	1	1
QY	713 PYSLSGNYWYPTNKNKGIGQDEKGISGVKVTLLKDENGNISITTTDENGKYOFDMJNSG	772	1	1	1
Db	587 S-----DSGDSASDSD-----SASDSASDSASDSASDS	615	1	1	1
QY	773 NYIVHDFKPSGQTQTTSDGDDDEQADAGEVHVITTDHDFSIDNGYVDDSDSDSD	832	1	1	1
Db	616 D-----SASDSDDNDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSD	658	1	1	1
QY	833 SPDSDDSD	892	1	1	1
Db	659 SPDSDDSD	718	1	1	1
QY	893 SPDSDDSD	952	1	1	1
Db	719 SPDSDDSD	778	1	1	1
QY	953 SPDSDDSD	1012	1	1	1
Db	779 SPDSDDSDA	838	1	1	1
QY	1013 SPDSDDSDSDSGSD	1055	1	1	1

[illegible]

```

: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 4749
: LENGTH: 466
: TYPE: PR1
: ORGANISM: Staphylococcus epidermidis
US-03-134-001C-4749

```

[illegible][illegible]

Db 310 IWRSMWMDNEVAFNNNGSGGDIDKPVVPEQ 340

RESULT 8

US-09-327-536-2

Sequence 2, Application US/09327536

Patent No. 635477

GENERAL INFORMATION:

APPLICANT: FISCHETTI, Vincent A.

APPLICANT: ROCHA, Claudia

TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A

TITLE OF INVENTION: STREPTOCOCCI

FILE REFERENCE: 02927-008

CURRENT APPLICATION NUMBER: US/09/327,536

CURRENT FILING DATE: 1999-06-08

PRIOR APPLICATION NUMBER: US 08/714,402

PRIOR FILING DATE: 1996-09-16

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 1161

TYPE: PRT

ORGANISM: SFBBP gene

US-09-327-536-2

Query Match

Best Local Similarity 18.8%; Pred. No. 9.9e-11;

Matches 244; Conservative 189; Mismatches 447; Indels 421; Gaps 54;

5.9%; Score 331.5; DB 4; Length 1161;

18 KSNKYAIRKFTVGTASIVYCATLP-----GLGHNEAKAEENSVQDYKDSNTDELSDS 71

3 QKNSYKTL-SPLSLTLGFLGLLVLGIGSVGVGHAEFR-----NGA 43

72 NDSSDEKKNNVNNNSINFTDNNQIIRKETNNYDGIKRSRDRSTETNVDENEATF 131

44 NKQSSFE-----IKKVDQNN----- 58

132 LQKTPDQTHLFEVEKES--SVESNSISDTAQ--OPSHITINRESVQTSNVEDSH 187

59 -KPLPGATSLTSKDKGTSVQSFTSNKGIIVDAQNLQPGTYTLKEEPAAPGVC----- 111

188 VSDRANSKIKASNTESGKEENTIEOPNKVKEDSTTSQPSGTTNIDEKISNDELINLPIN 247

112 -----KTSRTW-----TAVYENGTYTKLVENPN-----G 136

248 EYENKARPLSTASQPSIKRTVNOIAEAGSNVNHLLKVDQSTEEGYDSEGVIAHD 307

137 EIRKAGSKDVSSSL-----OLENPKMSVVSKEYG-----TEVSSGAAPFYRNA 181

308 AENLIYDTEFVDDKVS-----GDTMTVDIDKNTVPSDLTDSFTIPK-IRKNGEELIAT 361

182 A---YFKMSFELKOKKSEITINPGDTFVLQDRRLNPKGISOD--IPKIIYDSANSPALAI 236

362 GTYNKKNQIYTFDYYDKYENIKAHLLKLSYIDKSKVPPNNMTKLDEYKTAALSSVNTK 421

237 GKVAENHQLIYTFDYIAGDKVQLSAELSLPLENKEVLEN-----TSSINFSST 287

422 I-----TVEQRPNE-----NRANLQSMPTNIDTKNHYEQIYIINPLAYS 463

288 IGGQELTYKGTGNVLYGNESTKESNYITNGLSNVGSIESTNTEGEFVWVYVVPNRTN 347

464 AKETNVNISNG-----DEGSTIIDSTIIKVKVQDNOLPDSNRHYDSEVED 513

348 IPYATMNLMGCRARSNTSDLENDANTSALGELYVEVEGKELPSSGYV-----D 400

514 VT-----NDYIAQGNNDV-----NINEGN--IDSPYIIKVISKYPKKDDYTTIQQTVT 562

401 VTKTLRLRDITAGLNGFGOMTKRQRIDEGNNIYQNKAFIITKGTGTDQS-----GKPLV 453

563 MOTININEYGEFRAS-----YDNTIAFSSSSGGGCG-----DLPEKTYKAGD 606

454 VQSNLASFRGASEYAAFTPVGGNYVFNELIALSPKSGSGSGKSEFTKPSITVANLKRVAQ 513

QY 607 YWEDVDKGIQNTNDNEKPLSNLVLTLYPPDGTSSKSVR--TDEDGKYQFDGLKNGELTYK 664

Db 514 LRFKMSSTDNY-----PLPEAFELRSNGNSOKLEASNTQGEVHFKDLTSG--TYD 564

QY 665 ITFET--PEGI-----TPILKHSGTNPA-----LD 687

Db 565 L-YETKAPKGYQVTEKRLATVTDTRPAEEMVWYWGSPHSSVKVEANKVPIVNHKEITLT 623

QY 688 SEGNSVW-----VTINQO-----DDMTIDSG-----FYQTPKYS----- 716

Db 624 FSGKIMENDRPPDRPAKIQVOLLONGKMPNOIOETVKDKDMSHYHFKDLKRYAKNOEY 683

QY 717 -----LGANYW-----YDINKDGIQ-GDDE-KGISGVK-----TL 745

Db 684 KYSEEVAVNPBGYKVSYLGNIFPTREFEVEEQNNFLEGNMEIKQSGSKIIDEDTL 743

QY 746 KDENGNIISTTTDEN-----GKQFQNL-----NSG 772

Db 744 TSFGKKIKWKNDTAENRPOAIQVLYADGVAEGQTKRISGSGNWSFEFNKLYNCTG 803

QY 773 NYIVHFKPSGMDTQTTSDSGDDQDADGEEVHYITPDHDFSIDNGYDDESDSDSDSD 832

Db 804 NDIY-----SVKEYTVPYTGIVTYSAN-----DIINTKREVIYIOGPKLEIEETLPLESG 854

QY 833 SDSDSDSDSDS--SDS 885

Db 855 ASGTTVEEDSRPDTLSGLSDEGQSGDMTIEDSAHIKFKSRKIDGKELAGATWMLR 914

QY 886 DSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 931

Db 915 DSSKRTITWISDQVKKDFYLPKRYTFVETAPDGYEIAAIFTVVEGQVYVNGKAT 974

QY 932 DSDS--SDS 977

Db 975 KGDTHYVWDAKPKTKGSGQYIDIEKLPDQGHSGSTTEIEDSKSSDLIIIGGQGEVYDT 1034

QY 978 DS 1031

Db 1035 TEDQSGMTGHSSTTEIEDSKSSDVIIGGQGVETEDPOTGYGSGCKTEVENTKL 1094

QY 1032 -----DSDSDNDSDLCNSSDKSTKDKLPOTGAN 1059

Db 1095 VQSFHFDMKEPESSNEIIPKDKPKSNTSLPATGEN 1129

RESULT 9

US-08-714-402-2

Sequence 2, Application US/08714402

Patent No. 5910441

GENERAL INFORMATION:

APPLICANT: ROCHA, Claudia

APPLICANT: FISCHETTI, Vincent A.

TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING

TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,402

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 016921-097
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1112 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-714-402-2

Query Match 5.6%; Score 315.5; DB 2; Length 1112;
 Best Local Similarity 20.1%; Pred. No. 7.2e-10;
 Matches 181; Conservative 137; Mismatches 279; Indels 303; Gaps 43;

18 KSNKVAIRKFTVGTASIVIGATLLE-----GLGHNEAKAEENSVDVKDSDNDELSDS 71
 3 OKNSYKL-SELLSLTGFIILLVFIIGLVGVGHAEETR-----NGA 43
 72 NQSSDEKNDVYNNNSITDDNNOIKKEETNNYDGIKRSDEKSTETVNDENATP 131
 44 NKGSFE-----IKKVDNN----- 58
 132 LQKPDNTHLREEVEKSS--SVESNSSIDTAQ--QPSHTTINRESVQTSNDVEDSH 187
 59 -KLPKATSLTSGDKDGTSGVOSTSNDKGIYDAONIQPGTYTLKETATADGYP----- 111
 188 VSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQSGVTNIDEKISNODELLNLPIN 247
 112 -----KTSRTW-----FTVYENGTYKLV-----N 132
 248 EYENKARPLSTSAQPSIKRYVNOALAEQGS-NVNHILKVTD--QGITEGYDSE--G 301
 133 PYN-----GEIISKAGSKDVSSSLQLENNKMSVSKYGGKTEYVSSG 172
 302 VIKAHDAENLYDVFPEVDKVS-----GDTMTVDIDKMTVPSDLDSFTPK-INDNS 355
 173 AADFYRHHAYFKMSFELKDKDSEITNPDDTYLQDLRLNPGIGIOD--IPRITIDSA 230
 356 GEIATGTYDNKKKQITTYFTDYVYKXENIKAKHLKLSYIDSKSVPNNTKLDVEYKAL 415
 231 NSPLAIGKYHAENHQLTYFTDYIAGLDKYQLSAELSLFLENKYLEN-----TSI 281
 416 SSYVKTI-----TVEQRENE-----NRTANLQSMFTNIDTKNHYVEQTIYI 457
 282 SNEKSTIGGOEITYKGVNVLYGNESKESNYITNGLSNVGSIESYNTETGEFVWYVYV 341
 458 NPLRYSAKETNNVINSNG-----DEGSTIIDSTIIKVKYKVDNQLPNSRNIYD 507
 342 NPKRTINIPATYTNLMWGFGRKRSWTSLDENANTSSAELGELYVEVEGEKLESYGV-- 399
 508 YSEYEDVT-----NDVYALQNNNDV--NINFGN--IDSPYLIKYSKYDPKDDYTT 556
 400 -----DVTKLTLRTDITAGNGFQMTKRRIIDFGNNIIOKAFIITKTYGTDOS----- 448
 557 IQQVTVMQTTINETYGEFFRAS-----YDNTIAFTSSGGGOG-----DLPEPK 600
 449 -GKPLVVOSSMLASFRGSEYAAFTPVGANYVFOHEIALSPKSGSGSKSEFTKPSIYVAN 507
 601 TYKIGDYVMDVDKGIQNTDNNEKPLSNVLYLVITPDGSKSVR--TDDEKQKQFQGLK 658
 508 LKRVAAOLRFKKMSTDVN-----PLPEAFAELRSSNGNSOKLEASSNQGVEHFQDLT 559
 659 NGILTYITFT--PEGY-----TPTLKHSGTNPALDSEGSNV-----VTINGQDDMT 704
 560 SG-TYDL-YETKAPKGIQGVTEKLAITYVDTKPA--EEMVTWGSSSHSSKYVEANNEVT 614
 705 IDSGFYQTPKYSIGNYVWYDTNKGIGQDDEKIGSVKYLKDENG--NISTTTTDE 760

DB 615 I-VNHKETTFFS-GKKIWMENRDP-----QREAKIQVOLL-ONGKMPNQIQEVTKDN 664
 QY 761 NGKYQFDNLNSGNYIVHFDRKSGMTQTDTTSDGDEDDADQEEVHYVITTHDDFSIONGY 820
 DB 665 DMSYHFQDL-----PKTDAKQOEIKYISV--EEVNVDPGY 696

RESULT 10

US-08-904-263A-4
 Sequence 4, Application US/08904263A
 Patent No. 6015889
 GENERAL INFORMATION:
 APPLICANT: LINDAHL, GUNNAR
 APPLICANT: STAHAMMAR-CARLEWALM,
 APPLICANT: MARGARETHA
 APPLICANT: STENBERG, LARS
 TITLE OF INVENTION: PROTEIN RIB, A CELL SURFACE PROTEIN THAT
 TITLE OF INVENTION: CONTERS IMMUNITY TO MANY STRAINS OF THE GROUP B
 TITLE OF INVENTION: STREPTOCOCCUS: PROCESS FOR PURIFICATION OF THE PROTEIN,
 TITLE OF INVENTION: REAGENT KIT AND PHARMACEUTICAL COMPOSITION
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: FALLS CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/904,263A
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: SVENSSON, LEONARD R.
 REGISTRATION NUMBER: 30,330
 REFERENCE/DOCKET NUMBER: 552-119P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1231 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-904-263A-4

Query Match 5.3%; Score 301.5; DB 3; Length 1231;
 Best Local Similarity 19.0%; Pred. No. 4.9e-09;
 Matches 258; Conservative 198; Mismatches 494; Indels 409; Gaps 62;

3 NKKNNLTKKKPKANKKNAIRKFTVGTASIVIGATLLC----- 43
 8 NSYDTLOTQK-----RESIKKFKFGAASVLIGISFLGFTGQGFNISTDTVFAAEVI 59
 44 -----LGHNEAKAEEN-----SVQDVKDSNTD----- 65
 60 SGSAVTLNTMNTKRVQNGRAYIDLVDKNGKIDPLDLITLNSPLKQAYIROGNGFTQ 119
 120 PSELTIVGAASINTVYKTDGSPHTKPDGVDIINVSILTIYNSALRDKIDEVKKRAEDP 179
 106 NYDGIKRSFDR-----ESTVNDEN-----EATFLQK-----TPDNT 140
 180 KWD-----EGSRDKYILISDDIKTIDIDNNPKTQSDIANRITETVNTLEKILVRIPADKNDP 236
 141 HLTEEEVKESSSVESNSSIDTAQPSHTTINREESVQTSNDVEDSHVSDPANSKIKESN 200


```

QY 669 TPEGTPPLKHSCTPALDSEGSWWTINGODMTIDSGFYOTPKYSLGNVWDTNKD 728
Db 2690 -----ESG---GMDSTANEV-----EAVKVS--FPTLKFFPAG-----PGRT 2722
QY 729 GLOGDEKISGVKTVLKDENGNIISTTTDENGK-YOQDNL-----NSGNIVHFDKP- 781
Db 2723 VIDYNGERTLDGFKFL-ESGGMSTANEVAVKVSFPTLKFFPAGSGRNVIDNGER 2780
QY 782 --SGMTQTTTDSG-DDDEODADGEEVHTTIDHDFSIDNGYDDESDDSDSDS-DSDS 837
Db 2781 TLEGGKFFLESGLDHTENDVNG-----VYIEGYPTIVLYPGKKSESYYQGSRSLSLDF 2836
QY 838 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 897
Db 2837 DFIKENGDGAGDND-DLDLEALPDMEDDDQKAVKDELQDAGDDDLLELEAEER 2895
QY 898 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 957
Db 2896 DLEEDDDQKAVDELQDAGDDDLLELEAEERPMEDDDQKAVKDELQDAGDEGL 2955
/ 958 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1017
Db 2956 DLEAEERPLEEDDDQKAVRDELQDGAADDDLELDEETDELE-EGDDDEQKIQKDEL 3014
QY 1018 DSDSDSGS---DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1046
Db 3015 HEDVDGKALYEAEQKAEADADADAELEED 3046

RESULT 13
US-09-262-666-26
; Sequence 26, Application US/09262666
; Patent No. 6346244
; GENERAL INFORMATION:
; APPLICANT: HORT, Carsten Molland
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 63462440 No. 6346244disk of No. 6346244th America, Inc.
; STREET: 405 Lexington Avenue, 64th floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262, 666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,122
; FILING DATE: 11-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3880,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-9655
; TELEFAX: 212-867-0123
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-262-666-26

```

```

Query Match 5.0%; Score 282; DB 4; Length 3052;
Best Local Similarity 19.7%; Pred. No. 1.9e-07;
Matches 207; Conservative 147; Mismatches 418; Indels 280; Gaps 47;

QY 141 HTEEEVSESSVSSNSIDTAQPSHTTINRESVQTSQDVEDSHVDFAN-----SKI 196
Db 2129 HLMSELPDE-----DMDROPVAVLGKKE-----SDDLTKDKIKEFNKLEGTI 2175
QY 197 RESNTEGKEENTIEQPNKVKEDSTTSQSGYTNIDEXISNDEL--LNLPINEYENKA 253
Db 2176 KPHLMSQLPEWMDKOPKV-----LVGKDELSDKIVLESKAIESJB 2217
QY 254 RPLSTTSAPSTKRVTV-----NOLAQGSNV-----NHLI 285
Db 2218 KDFLKGDASPIYKSGEIEENDSSVFOLVGNFEEVAFDEKKNVFEFYAPWCHCKOLA 2277
QY 286 KVTDSITEGVDSGVKAKHDAENLIYDVPEYDOKVSGGTMVTDI-----KN 336
Db 2278 PIMDK-LGETYKDHENIILAKNFEEVAFD-----EKNVFEFYAPWCHCKQ 2324
QY 337 TVPSDLTDSFTIPKIKNSGELIATGY-DNKKOITYTFD-----YDKYENIK 386
Db 2325 LAP-----IMDKLGE-----TYKDHENIILAKNFEDVAFDEKKNVFEFYAPWC 2368
QY 387 AHLK-----LTSTYDKSKVPNNNTKLDVEYKTALSSVNTIYVYORP----- 429
Db 2369 GHCKOLAPIMDKLTGETYKDHENIILAKNFEEVAFDEKKNVFEFYAPWCHCKOLAPIMD 2428
QY 430 -----NENRANTLOSMFTNI---DTKNHTYQTIYINPLRSKAE----- 466
Db 2429 KLGETYKEHODIYIAKNFEEVAFDEKNVFE---FYAPWCHCKOLAPIMDKLPTIYRD 2485
QY 467 -TNVNISSGDEGSTIIDD-STIIKVKV---GNQNLPSDNRIYDSEYEDVNTDXYA 520
Db 2486 HENIYIAKNHDE---IVNDPKRDVLVLYAPWCHCKRLAP-----YQELA-DTYA 2533
QY 521 QLGNNNDVINFGNIDSP---YIIKVIS---KDPKNDDYTTIQ----- 558
Db 2534 ---NATSDVLIAMKDSANEVAVKVSFPTLKFFPASADRTVIDYNGERTLDGFKFL 2589
QY 559 QTVMTQTTINEY-----TGEFRTASYDNTIAFSTSSGOGGDLPEKTYKIGDYWA 609
Db 2590 ESGGMSTANEVAVKVSFPTLKFFPASADRTVI-----DYNERTL---DGK 2636
QY 610 EDVYDKGIQNT-NDNEKPLSNVLVTLYTPDGTSKSVFDEDEGKYQFDGLKNGLYTKIFE 668
Db 2637 KLESGGMSTANEVAVKVSFPTLKFFPASADRTVIDYNGERTLDGFKFL----- 2689
QY 669 TPEGTPPLKHSCTPALDSEGSWWTINGODMTIDSGFYOTPKYSLGNVWDTNKD 728
Db 2690 -----ESG---GMDSTANEV-----EAVKVS--FPTLKFFPAG-----PGRT 2722
QY 729 GLOGDEKISGVKTVLKDENGNIISTTTDENGK-YOQDNL-----NSGNIVHFDKP- 781
Db 2723 VIDYNGERTLDGFKFL-ESGGMSTANEVAVKVSFPTLKFFPAGSGRNVIDNGER 2780
QY 782 --SGMTQTTTDSG-DDDEODADGEEVHTTIDHDFSIDNGYDDESDDSDSDS-DSDS 837
Db 2781 TLEGGKFFLESGLDHTENDVNG-----VYIEGYPTIVLYPGKKSESYYQGSRSLSLDF 2836
QY 838 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 897
Db 2837 DFIKENGDGAGDND-DLDLEALPDMEDDDQKAVKDELQDAGDDDLLELEAEER 2895
QY 898 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 957
Db 2896 DLEEDDDQKAVDELQDAGDDDLLELEAEERPMEDDDQKAVKDELQDAGDEGL 2955
QY 958 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1017
Db 2956 DLEAEERPLEEDDDQKAVRDELQDGAADDDLELDEETDELE-EGDDDEQKIQKDEL 3014
QY 1018 DSDSDSGS---DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1046

```

Db 3015 HFDVDCALYEEAOEKAEEADADAELEDEED 3046

RESULT 14

US-09-134-001C-5080

Sequence 5080, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 5080

LENGTH: 3696

TYPE: PR

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5080

Query Match 4.9%; Score 279; DB 4; Length 3696;

Best Local Similarity 21.0%; Pred. No. 3.6e-07;

Matches 199; Conservative 134; Mismatches 351; Indels 262; Gaps 44;

QY 2 INKNNLLTKKKPIANK---SNKYAIRKFTVGTASIVIGATLLFGLGHNKAAKEENSQ 57

Db 1264 IONNDATFEEKEVANNLVNASQONVISKIDNATTNQIDIVSDGROSINATPDSIK 1323

QY 58 -----DVDSN-----TDELSDSNOSSEKENDVYINNQSINTDNNQII 99

Db 1324 RNAKNDIDIAADKKIKIQRINDATDEIDQAN-RKIEAKIEAKDNQFNSTDO--V 1379

QY 100 KKEETNNYDIEK-----RSEDRTESTVVDENEATFLOKTPODNTHLLEEVEKSSS 152

Db 1380 NEATNGINKENTTPATYKSEAR-QAVONKANEQNIHIONTD--ATNEKQEPAIN 1434

QY 153 VESSNSSIDTAQPSHTTINRESVQTSVDNEDSHVSDFANSKIKESNTESGKEENTEQ 212

Db 1435 RVSAELARVQKQ-----INAEHTQGVKTIKDAITSL--SRINAVYERESARNAIEQ 1486

QY 213 -----PNKVEDS-----TTSQPSGYTNIDKISNQD-----ELL 242

Db 1487 KATQOTQFINNNDATDEEKEVANNLVATKQKS-LDNINSLSSNNDENAKVAGINEIA 1545

QY 243 N-LPINEYENKARPLSTTSAPRSIKRVTVNOLAEOGSNVHLIKVTPDSITE-----G 295

Db 1546 NVLPATAVSKAKKDIDOKLAQINOIOTHOTATTEEKEA--AIOLANKSNEARTATON 1603

QY 296 YDSEGVIAKADAENLIYDVFEEVDKVKSGDTMTVDIK-----NTVPSDLTD-- 344

Db 1604 EHSNNVGAQA--KSGNHEIELVMPDAHKKSADAKO-SIDKNYNEQSNTINTP--DATDEE 1659

QY 345 -SFTPIKIDKSGELIATGTDNK-----NKQITTYTDDVVKENIKAKHLKLSYIDK 397

Db 1660 KOKALDKIK-----IADAGYKNVDAQOTNOVSDAKTEADITITNOQANV----- 1705

QY 398 SKVNNNTKLDVEYKKTALSSVNT--ITVE-----YORPENNR--TANLQSMFTNIDKN 448

Db 1706 AKKSARVELDSKEDFKROJINATPNATEEKEKODAIQLNKRDEVKHL-----INQDRD 1761

QY 449 HTVEQ-----TIYINPLRYSAKETNVNISGNGEGSTIIDSTIIKIVYKGDNON 498

Db 1762 NEVDQHNKIGLQLEETIHANFTPKRSDALQLEQ-----TKFISQTELLINNKNKATNBE 1813

QY 499 LPDSNRIDYSEEDYTVDDYAOGLGNNDVNVINGNISPIYIIKIKSYKPRNKDQTYIQ 558

Db 1814 KDEAKRLEISKNTTITNINOQOTNOVD-----NAKDNQMNIEATII 1856

QY 559 QVTWQTTINEYGEFRASVDNTIAFSTSGOGDLPPEKTYKIGDYWEVDK---- 614

Db 1857 PATTIKT-----DAKTA-----IDKKAQ 1875

QY 615 -----DGIONTNDNEKPLSNVLTYLPDGTSKSVRTDEGKQFQDLK-NGLYTKITFE 668

Db 1876 QVTIINNNDATDEEKAARLVEKAKIE--AKSNITNSPTEREVNAKTWGLD-KINNI 1932

QY 669 TPBEYTPFLKHSCTNPALDSGNSVWYTINGQODMTIDSGFYQPKSLG--NTVWYDTN 726

Db 1933 QPSTQYKT-----NAKOBINDXAOEOLQIINTPATEEKEOEATVRVNAQIAQIIONIN 1988

QY 727 KDGIQDDEGIGSGVKTLLKDNENGIISTTTDENGKYOFPDNL-----NSGNYVHFDPK 781

Db 1989 AHSQOEYNEKSTNSI-ATISVQPNVIAKPTAISLNLQEOANNOKTLGNGN----- 2039

QY 782 SGMQTQTTDSGDDEQAD-----GEEVHYTTTHHDEFSDIN 818

Db 2040 -----ATDDEKEAKOLVYQKLINEQIOTKIHSTODNOVDN 2074

RESULT 15

US-09-134-001C-4054

Sequence 4054, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 4054

LENGTH: 699

TYPE: PR

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4054

Query Match 4.9%; Score 278; DB 4; Length 699;

Best Local Similarity 21.1%; Pred. No. 4.8e-08;

Matches 152; Conservative 122; Mismatches 233; Indels 154; Gaps 32;

QY 15 IANSNKYAIRKFTVGTASIVIGATLLFGLGHNKAAKEENSVDVKDSNTDDELSDSNQ 74

Db 12 VKTRQNKYSIKKFSVAGASSILIALLPWGGSSQA-AEQQODKGVENSTTOSIGDEREK 70

QY 75 SSDEKNDVYINNQSINTD--NNQIYKKEETNNYDIEK-----RSEDRTESTVVDENE 128

Db 71 LSEOSTQONKRVNKSNDVSTENESLNEPKNMEDLIQOQKDSQNDKSSVVEQNKEN 130

QY 129 ATPLOKTPODNTHLLEEVEKSSS-----VESSNSSIDTAQOPS--HTTINRESVQ 178

Db 131 EAFYKHSHEEPQEOYVLEKHAENNOTLNSKAAQSHEVKTFRSOLDNTTAQOEDQK 190

QY 179 TSDNVEDSHVSDFANSKIKESNTESGKEENTIEOPNK-VKEDSTTSQPSGYTNIDKISN 237

Db 191 ENLSKQDTQ--SKKTDLLRATGQNSQDSOSTEEVNRKVKND--TQOVTAKNDDKV-- 244

QY 228 QDELINLPINE-----YENKARPLSTTSAPRSIKR-----VTVQGLAAEQ 277

Db 245 --ETFNLSKSEEPKVDQANP--TTDKDKSSKDKGSHDGLANLESNAVAATTKQSKQ 300

QY 278 GSNVNLKIVTPDSITE--GYDSEGVIAKADAENLIYDVFEEVDKVKSGDTMTV--DID 334

Db 301 VSEKNE--DQTNKSAKQOYKNNDEIILVHGFNGFTDIDINSVGLTHNYGGKMKMIQDLE 358

QY 335 KTVPSDLTDSFTIPKIDNSGEI--IATGTVD--NKNKQITTYTDDVVKYENI---- 385

```
Db 359 ENGVEYEASISAFSGSNYDRAVELYYIKGRVDYGAHAHAAGYGERYKGYEGVYKDK 418
QY 386 ---KAHLKLTSTYIDKSKVPNNNTKLDVEYKTAHSVNTITYEYORPENETANLQSMET 442
Db 419 PGOKIHL-----VGHSMGQOTIROLEELLRHGNPEEVEYQ----- 453
QY 443 NIDTKNHTVEQTIYINPLKYSKE-----TNVNISGNGDEGSTITDDSTIK--YKYG 494
Db 454 ---KOHGE---ISPLFOGHDNMVSSITTLGTPHNGTHASDLGNEAIVROLAYDVG 505
QY 495 DNQNLDPDSNRIDYSEY--EDVTNDYAOLGNNNDVNNINNGNIDSPYIIRKISK----- 546
Db 506 KMYGNKDSRYDFGLEHMGKOKPNESTIO-----YKRVONSKLMSK 548
QY 547 ---YDPNRDDYTTIOQVTMTOTI--NEYTGEFTASYDNTIAFSTSSGOGDLPPEK 600
Db 549 DSGLDLTRDGATDLNRKTSLNPNIVYKTYTGE---STHKTLA---CKOKADLMFL 599
QY 601 TYKI-GDIYWEVDYK-----DGIQNTNDNEKPLSNVLVLTLPDGTSKSVRTDEDGYOF 654
Db 600 PFTITGNLIGKAKEKEMRENDGLVSVISSQHPNOKYVEAT--DKNOKGVMYVTPTKHDW 657
QY 655 D 655
Db 658 D 658
```

Search completed: March 14, 2003, 13:01:18
Job time : 45.6129 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 13:00:30 ; Search time 16.3082 Seconds
(without alignments)
3086.325 Million cell updates/sec

Title: US-09-147-405B-15
Perfect score: 5646
Sequence: 1 MINKKNLTKRRKPIRANKSN.....FAGIGALLGKRKRKRKN 1092

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2-6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2-6/ptodata/2/pubpaa/PC7_NEW_PUB.pep:*
3: /cgn2-6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2-6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2-6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2-6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2-6/ptodata/2/pubpaa/PC7US_PUBCOMB.pep:*
8: /cgn2-6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2-6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2-6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2-6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2-6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2-6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2-6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2069	36.6	1349	10 US-09-815-242-5898	Sequence 5898, Ap
2	2069	36.6	1349	10 US-09-815-242-13137	Sequence 13137, A
3	2011	35.6	932	10 US-09-815-242-5578	Sequence 5578, Ap
4	2011	35.6	932	10 US-09-815-242-12438	Sequence 12438, A
5	1634	28.9	1021	10 US-09-815-242-5471	Sequence 5471, Ap
6	1634	28.9	1021	10 US-09-815-242-12544	Sequence 12544, A
7	1441.5	25.5	841	10 US-09-815-242-5779	Sequence 5779, Ap
8	1441.5	25.5	841	10 US-09-815-242-12751	Sequence 12751, A
9	1023	18.1	265	12 US-10-073-256-78	Sequence 78, Appl
10	865	15.3	1113	10 US-09-815-242-5836	Sequence 5836, Ap
11	865	15.3	1113	10 US-09-815-242-12769	Sequence 12769, A
12	607	10.8	2344	10 US-09-815-242-12713	Sequence 12713, A
13	500	8.9	1001	10 US-09-815-242-12686	Sequence 12686, A
14	491.5	8.7	1018	10 US-09-815-242-5797	Sequence 5797, Ap
15	491.5	8.7	1018	10 US-09-815-242-12838	Sequence 12838, A
16	477	8.4	767	10 US-09-815-242-5899	Sequence 5899, Ap
17	477	8.4	767	10 US-09-815-242-13140	Sequence 13140, A
18	458	8.1	978	10 US-09-815-242-5456	Sequence 5456, Ap
19	453	8.0	560	10 US-09-815-242-13057	Sequence 13057, A

20	448	7.9	502	10 US-09-815-242-5904	Sequence 5904, Ap
21	382.5	6.8	2368	10 US-09-815-242-5635	Sequence 5635, Ap
22	382.5	6.8	2368	10 US-09-815-242-12389	Sequence 12389, A
23	382.5	6.8	2478	10 US-09-815-242-5816	Sequence 5816, Ap
24	382.5	6.8	2478	10 US-09-815-242-12967	Sequence 12967, A
25	349	6.2	345	10 US-09-813-820-7	Sequence 7, Appl1
26	322.5	5.7	1031	10 US-09-815-242-10932	Sequence 10932, A
27	320	5.7	665	9 US-09-820-843A-107	Sequence 107, App
28	295.5	5.2	102	10 US-09-864-761-44209	Sequence 44209, A
29	287	5.1	1400	10 US-09-764-176-7	Sequence 7, Appl1
30	286.5	5.1	2150	9 US-10-135-322-17	Sequence 17, Appl1
31	286	5.1	1060	10 US-09-955-909-2	Sequence 2, Appl1
32	277.5	4.9	870	10 US-09-815-242-5493	Sequence 5493, Ap
33	277.5	4.9	870	10 US-09-815-242-12637	Sequence 4, Appl1
34	271	4.8	1115	9 US-10-153-273-2	Sequence 2, Appl1
35	267.5	4.7	2122	9 US-09-813-214A-9	Sequence 9, Appl1
36	264.5	4.7	489	10 US-09-876-889-350	Sequence 350, App
37	263.5	4.7	1770	10 US-09-841-132-444	Sequence 444, App
38	260	4.6	1183	9 US-09-870-759-45	Sequence 45, Appl1
39	259.5	4.6	1435	9 US-10-153-273-4	Sequence 179, Appl
40	257	4.6	1776	10 US-09-841-132-179	Sequence 34, Appl
41	254.5	4.5	1139	10 US-09-801-574-34	Sequence 3, Appl1
42	252.5	4.5	1786	9 US-09-742-096-3	Sequence 1835, App
43	250.5	4.4	821	10 US-09-841-132-195	Sequence 5835, Ap
44	247.5	4.4	2434	10 US-09-815-242-5835	Sequence 12996, A
45	247.5	4.4	6281	10 US-09-815-242-12996	

ALIGNMENTS

RESULT 1
US-09-815-242-5898
Sequence 5898, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA, 011A
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5898
LENGTH: 1349
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5898
Query Match 36.6%; Score 2069; DB 10; Length 1349;
Best Local Similarity 36.6%; Pred. No. 3.1e-92;

Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;

```

Qy 1 MINKN-NLTKKPIANKSKYARKTGTGTASIVGATLLEFGIGHAEAEKSVODV 59
Db 1 MLNENKATIRKQVNSRLKFSIRKTYGTASILVGTLLFGLGNEAKAESTNEL 60
Qy 60 KDSMTDELSDNOSSEKNDVNNOSINTDNOIILKEEENNDGEEKSEDETE 119
Db 61 NEAT-SDNOSSD-KVDMQOLNEDNTKNDNO-KEM-V3SQNETTS 105
Qy 120 STTVNENATELOKTPODNTHTLTEEVEKSSSVESNSIDTAQPSHTTINRESVOT 179
Db 106 NGNKLIEKES-VOSTGNKVEVSRAKSDQASPKSTMEDINTNQ-TLSNEALQ 157
Qy 180 SDNVEDSVSPANSKIKESSTESKEENTIEOPNKVKEDSTTSPSGTYNT-DEKISNO 238
Db 158 POLDENKSVN-VQPTNEENKVD-AKTESST-LNVKSPAIKSN 198
Qy 239 DELL-NLPIEYENKAPLSTSAQ-SIKRVNOLAEOGNSVNHILKVTQDSI 292
Db 199 DELVDNNSNNSNNENADILPKSTAPKRLNTRHRIAVOSSSTEAKVNDLITSNTLT 258
Qy 293 TEGYDSEGVIAKHAENLIYDFEVDKVKSGDTMTVDIDKNTV-PSDLTDSF 346
Db 259 VVADKNNKIYPAQOYLSLKSQIT-VDDKVKSGDYFTIKY-SDTVOYGLNPEDIKN- 313
Qy 347 TIPKTD-NSGEIITATGTYDNKKNKOITFTDYDKYENIKAHKLISYIDKSKVNNNT 405
Db 314 -IGDKIDPNNEGELTATAKHDANNLITFTDYDFRNSVQMGIVSYIMADTIP-VS 370
Qy 406 KLDVEKTAALSVNKTIVETVORPN-ENRPTANLOSMFTNIDTKNHT-VEOTI 455
Db 371 KNDVEFNNTIGTTTKTANIOYPIYVNEKNSISAFET-ETSHVGNKKNPGYKOTI 428
Qy 456 YINPLRYSAKETNNVNI-SGNGDEGSTIIDSTIIKVKYKGNQNPDSNRIVDYS- 509
Db 429 YVNPSENSLTNAKLKVOYHSSYPNNIGQINKDVIDIKIYQPKGYTL-NKGYDVATK 485
Qy 510 EYEDTNDYAO-LGNNNDVNIQGNIDSPYIIKIVSKIDPNKDDVTTOQTYTMOQT 566
Db 486 ETTDTN-QYLOKITTYGNNNSAVIDGNADSAVYVWVNTKFOYTNSPSTLWQMASTST 544
Qy 567 INEYGEFRFASDYNTIAFSTSSGOGDLPPEKTYKIGDYWVEDVDKGIQ- 618
Db 545 GN-KSVSTGNALGFTNNQSGAG-QEYKIGINYMWEDTNNNGVQLEKKGVG 594
Qy 619 NTN- 621
Db 595 NVTYVFDNNTWTKVGEAVTKEDGSYLLIPNLPNGDYRVEFSNLPKGYEYTPSKQNNBEL 654
Qy 622 DN-EKPLSNVLT- 634
Db 655 DSNGLSVITVNGKDNLSADGIYKPKYNLCDYVWEDTNKNGIDQODEKGISGVTYTLKD 714
Qy 635 TYPDG- 639
Db 715 ENGDLKVTITDADOKYFTDLHNGNKYVEFTTPEGYPTTIVSSGDIKESNGLTTGV 774
Qy 640 TSKSV- 644
Db 775 INGADNMTLDSGFYKTPKYNLGNVWEDTNKDKQDSFEKGISGTVTLKNENGEVLTQT 834
Qy 645 KIDDEGKIQFODLKNGLTYKITFEPEGTPPLKSHSGTRPALDSEGS- 692
Db 835 KTDKDKKYQFTGLENG-YKVEFEFTPSGYTPQVSGSGTDEGIDSNGSTTGVIKDKDNDT 893
Qy 693 VT-NGO- 700
Db 894 IDSGFYKPYNLGDYVWEDTNKNGVQDKDEKISGVTYTLKDENDKVLKTVTDENGKIQ 953
Qy 701 DDMTIDSIFYQTPK 714
Db 954 FTDLNNGTYKVEFEFTPSGYTPSTVSGNDEKDSNGLTTTGVIKDADNMTLDSGFYKTPK 1013

```

```

Qy 715 YSLGNVWYDNRKNDIGDDEKISGVKVTLKDENGNITSTTTDENGKYOEDNLNSGN 774
Db 1014 YSLGDYWYDNRKNDIGDDEKISGVKVTLKDENGNITSTTTDENGKYOEDNLNSGN 1073
Qy 775 IVHFDPSGMYTQTTSDGDEODADGEVHTTTDHDPSIDNGYVDESDDSDSDSD 834
Db 1074 KVIFEKPRAGLQTYTNTTEDD-KDADGGEVYTTTDDHDFLNDKRYFEED- 1123
Qy 835 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 894
Db 1124 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1181
Qy 895 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 954
Db 1182 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1241
Qy 955 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1014
Db 1242 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1300
Qy 1015 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1073
Db 1301 TTKDHNKAKALPETGSENGSNMATELFGLE 1332
Qy 1074 AGLCALL-GKRRNRK 1089
Db 1333 AALGSLLEGRKRNK 1349

RESULT 2
US-09-815-242-13137
Sequence 13137, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ECTTRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13137
LENGTH: 13137
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-13137

Query Match 36.6%; Score 2069; DB 10; Length 1349;
Best Local Similarity 36.6%; Pred. No. 3,1e-92;
Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;

```


[illegible]

	Matches	456:	Conservative	155:	Mismatches	334:	Indels	240:	Gaps	377:
QY	18	KSNKYAIRKFTVGTASIVAGIATLLEG-IGHNEKAKEANSVODVKDSDNTDELSDSNDSS	76							
Dd	5	KKKHAIIRKKSIGVASIVGLTGLGILLSREDASENSV-----TGSDASNN	52							
QY	77	DEEKDVINNNOISITDDONNOLIKKEELINNDIGIEKRSDRFESTTNDENPATLQLQTK	136							
Dd	53	EKSSEDSSSVNAPAKPTDVTNVSDRTTSTNTNG-----ETSAQNPAQ--OEIT	99							
QY	137	GD-NHLEEEK---ESSVSNESSSIDTAQPSPHTTIINREESVOTSDNVDSHVSPA	192							
Dd	100	QSASNTAITEETPVTEGATEATTAKOANPATTOSSNT--MAEELVNQTSNETAS-----	151							
QY	193	NSKIENSTESGEKENITEOPKNKVEDSTTSPSGSYTNIDEK--ISNDELINLPINEYE	250							
Dd	152	NDNTNVSSVNSPQNSTMNEWNSTTDISTEARFPPNNESAPOSTDASNN-D-VNQAVN---	207							
QY	251	NKARPLSTTSAOPSIKRYTVNOIAAE---QGSNNVHLIKYVDOSITEGYDOSEGYIKAHD	307							
Dd	208	-----TSA-PRMRASFSLAAVAADAPAAGDTIN--OLTGVKVT--IDSGTIVYPHQ	253							
QY	308	AENLYDYTFEVDOKVKGSDPTFTVIDDKNTVPUSDLTGSTIPKIXINDSGELLATGTYNOK	367							
Dd	254	AGYVLKNTGFSPNSAVXWDFKFIIVPEKLNINGVSTAKVPPIWGD-QVLANGLIDS-	311							
QY	368	NKOITYTFETDYDAKLENIKAHLLTSYIDSKSVPNNTKL-DVEYKTAL--SSVUKTIV	424							
Dd	312	DGNVITYTTFDYDMKENTVANTMTAYID---PENVYTKGVTILTLTGCTAFAKTVLI	367							
QY	425	EYORPENENTANLOSMTNIDTKNHTVEQTIVINPLRSAKETNV-----NISGNODE	477							
Dd	368	DYEKGQEHNLISIKETIDQIDKTNNTYROTIVYNP-----SGDNVLPALTIGNLIPNKS	422							
QY	478	GSTIIDSTIIIVYVVGDNQNLPSDRIRYDEEYEDTDDOVAQLGNNNDVANINF----G	533							
Dd	423	NALIDAKNTDIKVYR-DNANDLSSEYIVNPDEFEDVYNOVRISEPNAQYVERPTDD	481							
QY	534	NIDSPYIIKVIKSYRPNKDYYTIIQOTVMQTTINEYTEGF--RTASYDNITAFSTSGO	591							
Dd	482	QITTYIYVNVNHIDPASG-----DLARKFTFYGDNFIMRKSMNDMEVAFNNGSS	535							
QY	592	GGG-DLP--PEKTYRIGDY--VWEDVDKDGIONTDNENEKPLSNVLVTLTYPDGTSKSVRT	646							
Dd	536	GGGIDKPVPVEQDPBPGEIEPIPEDSDS-----PESDSCSDSN	574							
QY	647	DEDGKYQEDGLKNGLTXYKTFETPEGYTFLKHSGTNPALDEGNSVWYTINGOD-DMTI	705							
Dd	575	SDSG-----SDSGSDSTSDSDSDSDSDSDSDSDSDSDSDSDSDAS	607							
QY	706	DSGFTQTPKYSLGNYWDTKNDGIQGDDEKISGVKVTYLKENGILLSTTTDENGKQ	765							
Dd	608	DS-----DSTSDDSASDSASD-----SDSASDSASD	638							
QY	766	FDNLNSGNVIYHFDPSCGMTQTTTDSGDDDEODADAGEEVHYITDHDDFSJINGYYDES	825							
Dd	639	SBSASDS	697							
QY	826	DS	885							
Dd	698	DS	757							
QY	886	DS	945							
Dd	758	DS	817							
QY	946	DS	1005							
Dd	818	DS	877							
QY	1006	DSDSDSVSD	1046							
Dd	878	DSDSDASD	937							

	Query Match	25.5%	Score 1441.5	DB 10	Length 841	
	Best Local Similarity	32.9%	Pred. No. 2.6e-62			
	Matches 366	Conservative 158	Mismatches 294	Indels 293	Gaps 29	
QY	2	INKNLNLLTKRRKPIANKSNKYAIRKFTVGTSIVIGATLLFGLGHNEAKAEENSVDVAD	61			
DB	1	MNNKLTATNRKGMIPNRNLKFSIRKSYGTASIIIVGTLLFGLSGHEAKAE-----	52			
QY	62	SNTDELSDNSDQSSDEEKNDVINNOGSINTDDNNOIIKKEETNNVDJLEKSEDETEST	121			
DB	53	-HNGELNQSCKNEITAPSEBKTKKVDSSQLKDNO-----T	88			
QY	122	TNVDNEATFLOKTPQDNTHTLHEEVKESSEYSSNSIDTAAQPSHTINREESVOTSD	181			
DB	89	ATADQPKVT-----MSDAIVKETSNN-----QSPNATANOSTRTSMV	129			
QY	182	NVEDSHVSDAFANSKIKESNTESGKEENTIEQNNKVKEDBSTQPSGTYHIDEKISNDEL	241			
DB	130	TYNDSSSTTYSIEDTKSNLTOA-----KDVST-----	157			
QY	242	LNLPINENKARPLSTSAQPSIRKRVYNQAAE-EOGSNNVHLIKVT--DOSITEGY-D	297			
DB	158	-----PRTYTIKPRTLNRMAVNTAAQQGGINVADKHFNSIDALDAGHNV	204			
QY	298	DSEGVK--AHDAENLIDVTFEVDUDKYSQGTMTVDDIKNTVPSDLDSFTIPIKIDNS	355			
DB	205	QTTGTEWMAATSSDVLKLRKANTYIDDSVKEGTFEFKQYRPPSSVRLP.SOTONLYNAQ	264			
QY	356	GEIATGTYDMKNKQIITYTFTDYDVYKYEIRKHLKLTYSIDSKVEFNNTKLDVEKTA	415			
DB	265	GNIITAKGIYDSTNTTTFYTNVQOYTNVRSFEOVAARAKKNATTDVTAAKMEVTLGN	324			
QY	416	SSVKNITVEYORPENRNTALQSMFTNIDTKNHTVEQTYIIN-PLURSAKETNV-NISG	473			
DB	325	DYVESEIIVDY--GNKRAQPLISSTNYINNEDLSRNMATYANOPANTYTKOTFVNLTP	380			
QY	474	NGDEGSTIIDOSTIKYKVGVDGNL-----PDSNRIVDSEYEDVTNDYVQOLGNMNV	528			
DB	381	-----GYKNPNAKNFKIILEVDQNOFVDSFTPTYSKLKVDYQDFVI--YSDNKTATV	433			
QY	529	NINFGNIDS--PYIIKVI-----SKYDPKDDYTTIOQTVMQOTINETYGEFRFASYN	581			
DB	434	DLMKQOTSSNKQYIIQYVAYPDSSSTGDKIDY-----TLDPTKTKYSW---SNSYSN	483			
QY	582	TIATPSSGOGGGLDPEKTYKIGQYVEMEDVDKDIQNTDNEKPLSNLVLTLLTPDGIS	641			
DB	484	VNGSTTAGND-----OKKYNLGIYWEEDNKKDKQA--NEKISIGYVYLKDSNGKE	534			

```

RESULT 9
US-10-073-256-78
; Sequence 78, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
; APPLICANT: Kreiswirth, Barry N
; APPLICANT: Nadich, Steven M
; TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
; FILE REFERENCE: 19124.0002
; CURRENT APPLICATION NUMBER: US/10/073,256.
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 265
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-10-073-256-78

```

[illegible]


```

OY 147 VKESSVSSSSNIIIDAOOPSHITINRESVQTSNVEHSHVDEFSANSKI----- 196
Db 17 LRDGMAVASGFCKIDTAYN-----NVNGKVDLXLDADKTKNNLSQIGAAVGVGTFFYKNGA 70
OY 197 -KESNTEBEGKEBNTTEOP-NK-VKEDSTTSQPSGCTINDEKISNODELLNLINEYENKA 253
Db 71 DGVNVOGOMALNMTDKKRVNKITVADNTNNHLDGOFH-----GQRINDVYLNDAAT 122
OY 254 RPLSTTSAPISIKRVTVNQLAAEGOSNNHLIKVNDOSITEGYDGS-----EGVYKAND 307
Db 123 STIATVAGKRWK-ATTDLDGDKSQKYNFL--TSSHMKNKYSNGIMNTNLEGVITTPR 179
OY 308 AENLIYDVF-----EVDDKVKSGDTMV-----DIDKNTVPSDLTDSFTIPKI 351
Db 180 QADLIDDEVTKOPIPHKTIREFPDLEBGSADVIVOKEDEDEKETT-----TPTKV 231
OY 352 KDNSEIITATG--TYDKNN-----KPIITFTPDYVK-----YE 383
OY 232 DPTGDVVEREPTTEVTKNRPVDELYHFAPEVPEOGKHDEBPNI.PI DSTEEVPGKPGK 291
OY 384 NIKAHKLTSYID-----KSKVPNNNTKLDEYKTPALSSVNTIIVEYOR 428
Db 292 NPETGKVVTTPVDDVTKHGPKAGEBEVKEEPIPEKKR--BFNPDLKFGEEKVTQEG- 347
OY 429 PNEKRTALQSMFTNIDPKNHTVEGTIYINPLRYSAKENVNISSGDDGSGTIIIDSTII 488
Db 348 -----GKEKTTTPTTINPL-----TGEKV-BEG-PTTEVYKRPD 382
OY 489 KVVYKGDNONLPDSNRIDYSEY-----DVYNDYAOLGNNN-----DVNTN 531
Db 383 EITQFG-GEVPGQHK-----DEFDNLPIDIGEEVPGKGINFETGKAVTTPVDDVTKH 437
OY 532 FGNIDSPILIKVSYID-----PNKDDYT---TIQOTVMQTTINNEYTGFEKTA 577
Db 438 GPKAGEPEVTKEEIPEKKRREFNPDLKPEEEVUTQEGGTEKTTTPTTINPLTGE--- 493
OY 578 SYDNTIAETSSGOGOLPEPEKTKIGYVWEDVDKDIQNTNDEKLESLNVLVLTVP 637
Db 494 -----KVGEGFTTEKRPVBEITQFGEEVPO-CHKDEFPDNLPI----- 534
OY 638 DGTSKSVTDEBQKQFOLKNGLYTKITETPEEGVTPLLKSGTNPALDSEGNSVWVTI 697

```

```

GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12713
LENGTH: 2344
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12713

Query Match      10.8%; Score 607; DB 10; Length 2344;
Best Local Similarity 24.1%; Pzed. No. 1,1e-21;

```



```

      PRIOR APPLICATION NUMBER: 60/206,848
      PRIOR FILING DATE: 2000-05-23
      PRIOR APPLICATION NUMBER: 60/207,727
      PRIOR FILING DATE: 2000-05-26
      PRIOR APPLICATION NUMBER: 60/242,578
      PRIOR FILING DATE: 2000-10-23
      PRIOR APPLICATION NUMBER: 60/253,625
      PRIOR FILING DATE: 2000-11-27
      PRIOR APPLICATION NUMBER: 60/257,931
      PRIOR FILING DATE: 2000-12-22
      PRIOR APPLICATION NUMBER: 60/269,308
      PRIOR FILING DATE: 2001-02-16
      NUMBER OF SEQ ID NOS: 1410
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 12838
      LENGTH: 1018
      TYPE: PRF
      ORGANISM: Staphylococcus aureus
      US-09-815-242-12838

Query Match      8.7%: Score 491.5; DB 10; Length 1018;
Best Local Similarity 19.5%: Pred. No. 1.5e-16;
Matches 232; Conservative 191; Mismatches 490; Indels 275; Gaps 41;

QY      5 KNNLLTKKKRPLANSNKAIRKFRFYGTASIIYAGALLRGLGN-EAKAEENSVDQDKDSN 63
Db      2 KNNL-----RGIKRHKIIGAASVFLGMIIVMGQDKRAAASEQKTTTVEENG 49

QY      64 TDDELSDSDQSDDEKMDVINNNOSINTDDNNQIKKEETNNYDGIRKRESDPTSTTN 123
Db      50 -----NSATDKNSTETQTATTATVNHIEETQSINATVTEQPSN-----ATQ 89

QY      124 VDENEATFLQKTPQDNTHLTEEVKESSSVSSSSIDTAQPSHTTINREESVQTSNV 183
Db      90 VTTEAPRAVAQPCQAOPANIEYVEEYKAEAKPOY-----KETTSODNS 136

QY      184 EDSHVSDFRANSIKRSNTESGKEENTIDQPNKVKRDSSTTSQPSGVTINDEKISNDELIN 243
Db      137 GDQROVDLTPKKATONOVAETQVE--VAQPRASE----- 169

QY      244 LPINEYENKARPLSTTSAPSIKRYVTNQLAAEGSNVHLIKTIDQSTIEGDSSEGYI 303
Db      170 -----SKRYVRASADVAAKEASN-AKYETGTDVTSKYTVELGSI-EDHNNTNKV- 217

QY      304 KAHDAENIYDVTFEYDVKVSGDTMYVDIDKNVPSDLTDSFTLPIKIDNS-----GEI 358
Db      218 EPHAGQRAVLKYLKFEENGLHOGDYDEFELSNVNTVHGSTARKEPEIKNGSVYMATGEV 277

QY      359 IATGYDKNKNOITFTFDYVDKYENIKANHLTLTYSIDSKVRPNNTKJADVEYKALSS 418
Db      278 LEGG-----KIRYFTNIDEDKVDVAELBINLFIKDYQTQVGN-----QITTSYL 324

QY      419 N-----KTTVEYQRPENRNTANLQSMFTINIDTKNHTVEQTIYINPLRYSAKETVNI 473
Db      325 NEEOYSKELDVYKKGIGIYVYANLNGSIETPKKANRRESHVAFIR--NNGTTSVTVYG 382

QY      474 NDEGSSTIIDSTIIKYVK-VGDNONLPDSNRIV---DYSEYEDVYNDVYQOLG--NNN 526
Db      383 TLMKSNQNGNQPKVRIFEYELGNEDDIAS--VYANTDTQSKFKEVTSNMSGLNLDNNG 440

QY      527 DVNINFGNIDSEYIIKIVISKYDPNMD--DYTTIQQYVMTQITINEYTGEFRAASYNTIA 584
Db      441 SYSLIENIDKTYVVVHYDDEYVINGTDEVDERT-QVNGHPEOLYKYYIDRGYGLTWONGV 499

QY      585 FSTSSGQGGDLPPEKTYIGDYVMEVDVKGQ--IONTNDNEKPLSNVLTITLYPQGTSK 642
Db      500 LYSNKAANG-----KNGBIIONK-----FEYKEDTIR 528

QY      643 SVRTDEDGKYQFQDGLKNGLYTKRITFETPGYTPTLKHSQTNPALDSEGSV--WVTJNG 699
Db      529 ERTL---GGYD---KNLVT---TVE--BEYOSSTLIDIDYHAIADGGGYYVGDIETIEE 756

QY      700 QDDMTIDSFYQTPKSLGNYVMYDINKNGIQGDDKKGISGVKVTILKDNNGINLIISTTTD 759

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01 ; Search time 22.1792 Seconds
(without alignments)
4733.207 Million cell updates/sec

Title: US-09-147-405B-15

Sequence: 1 MINKNNLLTKKKPIANKSN.....FAGLGALLGRRRKNKKN 1092

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

1: PIR-73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5646	100.0	1092	T30214	fibrinogen-binding
2	2698.5	47.8	1166	T28680	fibrinogen-binding
3	2589	45.9	1141	E89624	hypothetical prote
4	2234	39.6	1385	D89824	hypothetical prote
5	2093.5	37.1	953	C89824	hypothetical prote
6	1880	33.3	1315	T28679	fibrinogen-binding
7	1652	29.3	989	D89852	fibrinogen-binding
8	1628	28.8	933	S41539	fibrinogen-binding
9	1536	27.2	877	F90070	Clumping factor B
10	785	13.9	882	AG1671	probable peptidogly
11	782.5	13.9	903	AG1299	probable peptidogly
12	624.5	11.1	2271	F90073	hypothetical prote
13	554	9.8	406	S38170	SrB40 protein - ye
14	554	9.8	3394	T18501	hypothetical prote
15	549	9.7	2570	T17451	fimbriae-associate
16	525.5	9.3	1038	H90053	hypothetical prote
17	520	9.2	940	S19702	fibrinogen-binding
18	502.5	8.9	961	G90053	hypothetical prote
19	487.5	8.6	1018	A32192	fibrinogen-binding
20	471	8.3	334	A54138	acidic repetitive
21	471	8.3	640	A54502	S antigen precursor
22	452	8.0	2910	T28156	DNA-directed RNA p
23	442.5	7.8	1999	AB2018	hypothetical prote
24	435	7.7	1192	A71623	probable secreted
25	430.5	7.6	4550	T18440	hypothetical prote
26	391.5	6.9	1217	S52714	sericinib - silkw
27	388.5	6.9	695	S27390	calcium-binding pr
28	381.5	6.8	955	T18435	hypothetical prote
29	379.5	6.7	3724	T18427	hypothetical prote

30	371.5	6.6	1305	2	T00670	hypothetical prote
31	370	6.6	1063	2	D86731	lustrin A - Califo
32	368	6.5	1664	2	T18262	fmbb protein (limp
33	366	6.5	3844	2	T18402	hypothetical prote
34	364.5	6.5	792	2	T42963	hypothetical prote
35	363.5	6.4	1360	2	T18403	cell wall surface
36	362	6.4	1193	2	G71605	serine/threonine-s
37	351.5	6.2	1428	2	T08852	probable membrane
38	345.5	6.1	2481	2	D90011	hypothetical prote
39	343.5	6.1	1658	2	S55101	hypothetical prote
40	338	6.0	4776	2	E95206	agglutinin-like ad
41	336	6.0	2485	1	H71621	
42	335.5	5.9	4910	2	S64942	
43	333.5	5.9	1072	2	A86827	
44	333	5.9	797	2	A36811	
45	331.5	5.9	1419	2	T30531	

ALIGNMENTS

RESULT 1

T30214
Fibrinogen-binding protein - Staphylococcus epidermidis

C/Species: Staphylococcus epidermidis

C/Date: 22-Oct-1999 #sequence_rev15ion 22-Oct-1999 #text_change 11-May-2000

C/Accession: T30214

R/Nilsson, M.; Frykberg, L.; Floock, J.L.; Pel, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998

A/Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.

A/Reference number: Z20781; MID:98261511; PMID:9596732

A/Accession: T30214

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 11092 <NID>

A/Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CA176638.1

Query Match 100.0%; Score 5646; DB 2; Length 1092;

Best Local Similarity 100.0%; Pred. No. 2, 9e-211;

Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MINKNNLLTKKKPIANKSNKYAIRKFTVGASIVIGATLLFGIGHNKAKEENSVDYK	60
DB	1	MINKNNLLTKKKPIANKSNKYAIRKFTVGASIVIGATLLFGIGHNKAKEENSVDYK	60
QY	61	DSNTDDELSDNDSODEKNDVYNNOSINTFDDNNQIKKEETNNYDGIRKSEDPRES	120
DB	61	DSNTDDELSDNDSODEKNDVYNNOSINTFDDNNQIKKEETNNYDGIRKSEDPRES	120
QY	121	TTTNDENATFLQKTPQDNTLHTEFEVKESSSVSSNSIDTAAQPSHTTINRESYOTS	180
DB	121	TTTNDENATFLQKTPQDNTLHTEFEVKESSSVSSNSIDTAAQPSHTTINRESYOTS	180
QY	181	DNVEDSHVSDFRANSKIKESNTSGKEENTIEQPNKVKEDSTTSPSGTYNIDEXISNO	240
DB	181	DNVEDSHVSDFRANSKIKESNTSGKEENTIEQPNKVKEDSTTSPSGTYNIDEXISNO	240
QY	241	LLNPINIEYENKARPLSTTSQAQPSIKRTVVNOALAEQSSNNHLIKYDOSITTEYDSE	300
DB	241	LLNPINIEYENKARPLSTTSQAQPSIKRTVVNOALAEQSSNNHLIKYDOSITTEYDSE	300
QY	301	GVTKAHDAENLIYDVFEEVDKVKSGDTMTVDIDKNTVPSDLTDSFTTIPKIKDNGEIIA	360
DB	301	GVTKAHDAENLIYDVFEEVDKVKSGDTMTVDIDKNTVPSDLTDSFTTIPKIKDNGEIIA	360
QY	361	TGTVDNKKKQITFTTVDYDKYENIKAHLKLTYSIDKSKVPNNNTKLDEYKTAALSVNK	420
DB	361	TGTVDNKKKQITFTTVDYDKYENIKAHLKLTYSIDKSKVPNNNTKLDEYKTAALSVNK	420
QY	421	TTTTVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETVNNISGNDSEST	480
DB	421	TTTTVEYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETVNNISGNDSEST	480


```

Db      61  NATT-----SASDNOSD--KYDMOOLNQBEDNTKNDQ-----KEVSSQONE 102
Oy      120 STTNVDN-EAFTLOKTPQDNHILTEEVKSSVSSSSSIDFAKQPSHTITNPEAQ 178
Db      103 TTSNGKRSLEKSSVOSTGKNKVESTAKSDEQASPKSTVEDNTKQ-----TISNOEGIQ 157
Oy      179 TSDNVEDSHVDFANSKIKESNTESGKEENTIEOPN-KKEDSTTSQPSGYNIPEKISN 237
Db      158 -PDLENKSVVN-----VQPTNEENKVKADKTESITLANKSAIKS-----NAELVYN 205
Oy      238 ODELNLPIINEYENKARPLISTSAQPSIKRVYNOLAAEQS-----NVNHLIKVTDQST 293
Db      206 NSNSNN-----ENNADILPKSTAPKSLNTRMMAAIOQPSNDSKKNVNDLITSMITLV 259
Oy      294 EGYDSEGVIAKHAENLIYDTFEVDOKVSGDTMTVDIDKNV-----PSDLTSDT 347
Db      260 VDADSKSTIVPQDYLKSLQIT--VDKVKSGDYFTIKY--SDTYVYGLANEDIKN--- 313
Oy      348 IPKIRD-NSGELIATGYDNKNKQITFTDYVDKYENIKAKHLKLSYIDKSKVFNNTK 406
Db      314 IGDIDPNNGEITAKAKHTANNTLITFTDYVDFNSYKMGIN SYVADDTITVD-K 371
Oy      407 LDVEYKTAALSVNKTIVYEQRP--NENFTANLQSMFTNIDTKNH-----VEQTY 456
Db      372 KQVPSVITIGNQITTTADITPAYKEDMNSIGSAFT--ETVSHVGNVEDPGYNNQYV 429
Oy      457 INPLKSAKETVINISGNDEST-----IIDSTIIKYKVDNONLPSNRIYV--SE 510
Db      430 VMPKMDIKGAKLKVAHPKPYPTINGOINQVNTIKIRVEGYTL--NKGYVNTND 486
Oy      511 YEDVYND--DYAQLGNNDVNINFGNIDSPYIYIKISKYDPKDDYTTIOQVNTMOTIN 568
Db      487 LVDVIDEFKNKNTYGSNOSVNLDPQDITSAVYVWNTKFOYNNSSPILVQMATLSSIGN 546
Oy      569 EYTGEPFRASTONTAETSSGQGGDLPPEKTYKIGDYWEDVEKDGIO----- 618
Db      547 ----KSVTGNALGFTNNOSGAG--GEVYKIGNVWEDTKNGVQELGKGVGVN 596
Oy      619 ----NM----- 621
Db      597 TYTVVDNNTNFKVGAIVTKEDGSYLILPLNPGDYVEFSNLPKGYEYTPSKOGNEELDS 656
Oy      622 ----DN-----EKPLSNVLTL--- 634
Db      657 NGLSSVITVNGKDNLSADLGIKPKRYNLGDYWEDTKNGIDQDEKIGSGVTVLKDEN 716
Oy      635 ----TYPDG----- 639
Db      717 GNVLKTVTTDADGKYKFTDLDNGNFKVEFTTPEGYTPPTVTSGSDIEKDSNGLTTGVIN 776
Oy      640 ----TSKSY-----RT 646
Db      777 GADNNTLDSGFPIKTKYKYLGNVWEDTKNGKODSTEGKISGVTVLKNGENGEVLQTTKT 836
Oy      647 DEDGKYQDGLKNGLTLYKITEETPEGYPTLKHSSTNPAIDSEGS----- 692
Db      837 DKDKGYQTGLENG--TYKVEFETPSGYTPTVOGSGTDEGIDBSNGISTGVKDKKNDITD 895
Oy      693 ----VW-----VTI-----NGO----- 700
Db      896 SGFYKPYNLGDYWEDTKNGKNGVODKDEKIGSGVTVLKDENDKVLKTVTTDENKGYFT 955
Oy      701 ----DQMTIDSGFIQTPKYS 716
Db      956 DLNNGTYVEFETPSGYPTVTSGNDETEKDSNGLTTGVKADANMTLDSGFYKTPKYS 1015
Oy      717 LGNYWYDTNKGIGDDEKIGSGVTKLKDENGNIISTTTDENGKYOPNLNGNVIY 776
Db      1016 LGDYWYTSNKGDKQDSTEGKIKDYKVLKLNKGEVIGTITTDENKGRFENLDSGKRV 1075
Oy      777 HFDRKSGMTQTTTSGDDDEQADAEVHVTITDHDPSIDNGYVDE-SDSDSDSDSDS 835

```

```

Db      1076 IFEKPTGLTQTCTNTTEDD-KDADGGEVDVYTTIDHDDFTLDNGYIEERTSDSDSDSDS 1134
Oy      836 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 895
Db      1135 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1194
Oy      896 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 955
Db      1195 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1254
Oy      956 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1015
Db      1255 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1309
Oy      1016 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1064
Db      1310 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1359
Oy      1065 KGTLLGTLFAGIGALLL-GKRRKNR 1089
Db      1360 NATLFGILFALGSLLLFGRKKQNK 1385

RESULT 5
C89824
hypotheetical protein sdrc [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89824
R:Kusuda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: C89824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-953 <KUR>
A:Cross-references: GB:BA000018; PID:913700453; PIDN:BA041750.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrc

Query Match 37.1%; Score 2093.5; DB 2; Length 953;
Best Local Similarity 43.8%; Pred. No. 5.9e-74;
Matches 488; Conservative 140; Mismatches 298; Indels 189; Gaps 28;

Oy      2  INKKNLLTKKKPIANKSKNYAIRFTVGTASIVGATLLGIGHNEAKAENSVDYKD 61
Db      1  MNKKTATNRKGMIPNRKLSIRKYSVGTASILVGTLLIFGLSGHEKAAE----- 52
Oy      62 SNTDELSDNSDQSDDEKNDVYNNNOSINTDNNQIIEKETNYVDIEKRSDETEST 121
Db      53 -HTNGELNQSKNETAPSEN-----KTT 74
Oy      122 TNVDENEAFTLOKTPQDNHILTEEVKSSVSSSSSIDFAKQPSHTITNPEAQVSD 181
Db      75 EKVDSROL-----KDNQTA-----TADQPKVY----- 97
Oy      182 NVEDSHVDFANSKIKESNTESGKEENTIEOPNKKEDSTTSQPSGYNIPEKISN 238
Db      98 -----KSDATVKEETSSNMOS-----PQNTASOSTTQTSNVTNKSSTTYSE 142
Oy      239 DELNLPIINEYEN-KARPLSTSAQPSIKRVYNOLAA-EGGSNVNHLIKVTDQST 292
Db      143 TDKSNL--TQAKNNTSTTKTTTIKORALNRNAVTVAAPOGGINVNDKVHTNIDIAIDK 200
Oy      293 -----TEGYDSEGVIAKHAENLIYDTFEVDOKVSGDTMTVDIDKNVPSDLT 343
Db      201 GHVKKTGTEFMAITSSDVLK-----LKANYTIDSVAKEGDFFPKYQGYFRAPGSVR 252
Oy      344 DSFTIPKIKDNGELIATGYDNKNKQITFTDYVDKYENIKAKHLKLSYIDKSKVFN 403

```



```

Db 253 LPSQONLNAGNITAKITVDSKNTTFTFTNYVDQNTNVSQSFQVAFARREANATD 312
Qy 404 NTKLDEYKTAALSVNKTITVEQRPENRNTANLQSMFNIDRKHTEQITVIN-PLRY 462
Db 313 KTAAYKEVTLGHDYTSKDIYDV---GNQKQGLISTSTVNNEDLSRMNVTYVNOPKRT 369
Qy 463 SAKETNV-NISGNDGSGTIIIDSTIIKYKYGDNONLSDNRITVYSEEDVND-DVA 520
Db 370 YKRETFVTMLT-----GYKFNPDANKFKIYEVDQNOFPVDS-FTPTDSKLKLDTGQFDVY 423
Qy 521 QGNNNDVINFGN----IDSPILIKISKYDQNKDDYTTIIOOTVMOQTINYTEGFR 576
Db 424 YSNDKNTATVDLNGSSSSDKOTIIQOVAVYPDNSSTONGKIDYLTLETQNGKSWSN---- 479
Qy 577 ASYDNTIAFSTSGOGGLPPEKTKIGDYVEDVDKDIQONTNENKPLSNVLTLY 636
Db 480 -SYSVNSSTANGD-----QKYNLGDYVEDVINKGKQDA--NEKGIRKVVYILKD 529
Qy 637 PDGTS-KSVRTDDEGKYOFDGLKNGLYKITEFTEPGYPTLKHSGTNPALDEGNSVMY 695
Db 530 SNGKELDRFTDENGKYOFTGLSNG--TYSEFSTPAGYPTTANAGTDADVDGLTTTG 588
Qy 696 TINGODDMTIDSGEYOTPKYSLGNVYVDPNKGICGDEKISGYKVTLKDENGNISF 755
Db 589 VIKDADNMTLDSGFYKTPKYSLSGDYVYDKNKGKODSTEGIKGVKVTLOKNGEYIGT 648
Qy 756 TTTDENGKYOFTDLNNGNIVHFDPKPSGMTQTITDSDGDEODADEEYVHTITDHDPS 815
Db 649 TETDEGKTRFDLSDGKRVIFEKPAQLGTQNTTDD-KDADGEVYVTTIDHDFT 707
Qy 816 IDNGYVDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db 708 LNGYVEEET-SPSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 766
Qy 876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db 767 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 826
Qy 936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db 827 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 886
Qy 996 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1055
Db 887 DSDSDSDAGKHRTPKPMSTV-----KDQHKTKALPE 918
Qy 1056 TGANEDYSGKTLTGTLFGLGALLL-GKRRKRX 1089
Db 919 TGSENNNSNNGTLFGSLFALGSLLEGRKKONK 953

RESULT 6
T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28679
R:Accession, E: McCrear, K.; Ni Eldh, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of St
A:Reference number: 220510; M0ID:99098700; PMID:9884231
A:Accession: T28679
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CA06651.1
C:Genetics:
A:Gene: sdid
Query Match 33.3%; Score 1880; DB 2; Length 1315;
Best Local Similarity 36.0%; Pred. No. 1.6e-65;
Matches 498; Conservative 177; Mismatches 346; Indels 362; Gaps 46;

```

```

Qy 1 MINKN-NLTKKKPIANKSNKAIKRTVGTASTIYGAATLFGILGNKAKAENSVDVY 59
Db 1 MLNREKRTAITRGKAVSNLNKFSIRKYVGTASILVGLTFLGLDQEKKAESTNKL 60
Qy 60 KOSNTDELDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 119
Db 61 NEAT-----SASDQSSD--KYDMOQLOEDNTKNDQ--KEM-----VSSQNETTS 105
Qy 120 STTNVDENATFLQKTPRODNTLJEEVKESSSESSSSSIDTAQPSHTTINRESVOT 179
Db 106 NGKLLIEKS--VQSTGNKVEVSTAKSPDQAPKSTNEDLNTKQ-----TISNOALO- 157
Qy 180 SDVVEDSHVSDFAANSKIKESNTESGKEENTIEOPNKVEDSTISQPSGTYNT-DEKISQ 238
Db 158 PDLOEKRSYVN-----VOPTNEENKRYD-----AKTESTT-----LNVKSDAISK 198
Qy 239 DELL--NLPINEYENKARELSTSAOP-----SIRKRVNOLAEOGSNVHLKVDQSI 292
Db 199 DETLVNNSNNSNNENNADITLPKSTAPKRLNTMRJAAVOPSTKAKNVNDLITSNTLT 258
Qy 293 TEGYDSEGVIAKHADENLIDYTFEVDKVKSGDTMTVDIDKNTV-----PSDLTSP 346
Db 259 VVDADKNNKIVPAQDYLSLKSQIT--VDQVKSGDYFTIKY--SDYQVYGLNPDIKN-- 313
Qy 347 TTPKID-NSGELIATGYDNKKQITTTTVDYKYEKIKAKLSTYIDKSKVPNNNT 405
Db 314 -IDIDIPNNGETIAAKHDTANLITFTTVDVDRNSVQOMINSIYMDADTIP--VS 370
Qy 406 KLVEYKTAALSVNKTITVEYORPN--ENRTANLQSMFNIDRKHTEQITVIN-PLRY 455
Db 371 KNDYERVNTGNTTTTNTTANIOYPDYVNEKNSIGSFT--EYSHVGNKENGKYKOTI 428
Qy 456 YINPLRSAKETVNI-----SGNGDESGTIIIDSTIIKYKYGDNONLSDNRITVY 509
Db 429 YVNPSESLNNAKLQVAYHSSYPNNIGQINKVDITKIKYQVKGYL--NGGYDNTK 485
Qy 510 EYEDYTRDDYAO--LGNNDVINENIDSPILIKYIKYDPRNKDYYTIOQTYMOT 566
Db 486 ELTDVYN-OYLOKITGYDNNNSAVIDEFNMASSAVVYNTFOYTNSPPLVGMATLST 544
Qy 567 INEYTGFTASYDNMTIAFTSSGOGGLPPEKTKIGDYVEDVDKDIQ----- 618
Db 545 GN-----KVSYGNALGFTNNOSGAG--GEVYIGNRYVEDTKNGVOELGKGVG 594
Qy 619 -----NTN----- 621
Db 595 NNTVTVFEDNNTFKYGEAVTKEDGSYLIPLPNQDYVEPSNLPKGYEVRPSKQNNEL 654
Qy 622 -----DN-----EKPLSNVLTLY 636
Db 655 DSNGLSSVITVNGKDNLSADLGIYKPKYNLGDYVWEDTNKNGIODDEKISGTYVLKD 714
Qy 637 PDG-TSKSVRTDDEGKYOFDGLKNGLYKITEFTEPGYPTLKHSGTNPALDEGNSVMY 695
Db 715 ENGVLTCTVTTDADGKTKKFDLDNG-NYKVEFTPEBTPTTYTYSGSDIKKDSGLTTTG 773
Qy 696 TINGODDMTIDSGEYOTPKYSLGNVYVDPNKGICGDEKISGYKVTLKDENGNISF 755
Db 774 VINGADNMTLDSGFYKTPKYSLSGDYVYDKNKGKODSTEGIKGVKVTLOKNGEYIGT 833
Qy 756 TTTDENGKYOFTDLNNGNIVHFDPKPSGMTQTITDSDGDEODADEEYVHTITDHDPS 815
Db 834 TKTDKDGKYOFTGLENGTYVEFEETPSGYPTVOGSGTDEGIDSNSTGTGVIKDKDNDT 893
Qy 816 IDNGYVDE-----SDSDSDSDSD----- 836
Db 894 IDSGFYKPTYNLGDYVEDVINKNGVQDKDEKIGSGVYVTLKDENKYKLTVTYDENGKYO 933
Qy 837 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS--DS----- 861
Db 954 FTDLNNGTYKVEFEETPSGYPTVTSNGDREKNSNGLITTGVIKADANMMLDSGFYKTPK 1013

```


A;Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Query Match	28.88	Score 1628	DB 2	Length 933

Matches 420; Conservative 153; Mismatches 313; Indels 230; Gaps 35;

QY .18 KSNKYAIRKFTVGTA SIVIGATLLFG-LGHNEAKAEENSVDVKDSNTDDELSDSNDQS 76

Oy	.18	KSNVATRKFTYOTASIVIGATLTF--LGHNARKAEENSVDVKDSNTDELSDSNOSS	76
Dd	5	KKEKHAIKRSKISIVASVLVGTLIGFGLLSRKADASENSV-----TQSDASN	52
Oy	77	DEENVDYNNOSSINTDDNNQIITKKEETNNYOICEKREDETRFEESTYNDENEAFLOKTP	136
Dd	53	ESKSNDSSVSAPAKPDIDTVN-----SDPTTSKDTN--NGEFSVANQP	93
Oy	137	ODNHLEEEVEKKSESSVESSENSSIDTAOOPSHATTINREESVOTSONDVSHVSDFANSKI	196
Dd	94	-----AQGETTOSSTNNATEETPYTGGAATTTTNQANTPATPGS-SNTVAEELVN---	143
Oy	197	KESNTEGCKEENTIEOPNKVKEDSTTSQSQSYGNIDEKISINODELLN-LPINEYKARP	255
Dd	144	QTSMETTFNDNTV-----SSVNSQNSTNEANSTJODUSTEATPSN--NESAP	191
Oy	256	LSTNSAOPSIKRTYVN-----OLAAEQSNVHLILKYDVQ--SITEEGSEGYIK	304
Dd	192	OSTASKKDYVNOAVNTSAPRMRAPFLAVALAADAPAAGDITLNQNLNTVTGI-DSGTTY	250
Oy	305	AHDENILTYDETEVDKDKSGDTPMDVIDKNTVSDLTSTFKIKIDNSGEIIATGT	364
Dd	251	PHOAGVKLYKGFSVPBSAVKGFPIITPKELNLCGVSTKRVPPIMAG-QYLANGVI	309
Oy	365	DNKNQOTYTTFTEVDYEENIKALHLKTSYIDSKPVNNTKL-OVEKTKALSS--VAKT	421
Dd	310	DS-DSNYITFTDVTNNKDVKATFLMPAID---PEVKKTGVNVLATIGGSTAKT	364
Oy	422	ITVEYORPENRFANLOSMTNIDTKNHTVEQTYYINLP--RYSAKETNVINSNGDEGS	479
Dd	365	VLVYEEKYKGYNLISIGTIDQIDKNTNNTRYROIYVNPBGDVIAPVLTGLMKPTDSNA	424
Oy	480	TIIDSTRIKYYKYGQNQLPDSNRIFYDSEYEDVNDYAOJCANNPNVINFG----NI	535
Dd	485	LIDQONTSIKYKKYDMADLSESYFV-NPENFDYTNVNITFPNPOKYKFENIPDDOI	483
Oy	536	DSPLYIKVISKYDEN-KDYTTYTIOQTVTMOTTINETGE--FRITASDYNTIAFSYSGOG	592
Dd	484	TTPYIVVYNGHIDNSSGD-----LALSTLYGNSNIIMRSMSMEVAFFNNGSSG	536
Oy	593	OGDLPERPTYIGDYWVEDVDKDGIONTNEKPLENVLTLYPDGTSKSVRTDECKY	652
Dd	537	DG-----IDKPYVP----BDP-----DEPEI	554
Oy	653	QFDGLKNGLTKITFETPEGYTPLLKHSCTNALDSEGNSVWVTINGODMTIDSGFYOT	712
Dd	555	E-----PIPEDSDDPCSDSGSDS-----NDSGSDSGSDST	586
Oy	713	PXYSLGNTVMYDTRKKGDIQGDDDEKGISGVVTLKDENGMIISTTTTDENGKQKPONLSG	772
Dd	587	S-----DSGSDASBSD-----SASDSASDSASDS	615
Oy	773	NVIYHFDPGSMGTTTDDSGDDDEQADAGEHVHTITJDHDFSIDNGYDDESDBSDSD	832
Dd	616	D-----SASDSDBSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSD	658
Oy	833	SD	892
Dd	659	SD	718
Oy	893	SD	952
Dd	719	SD	778
Oy	953	SDSV	1012
Dd	779	SDSDSDSDSDSDSDSDSDSDSDSDSDSESDSDSESDSDSDSDSDSDSDSDSDSDSA	838

```

Qy 1013 SDSPSDSDSPGSDSPSDSDSDSDSDSDCN-----SDKSTYDKLPD 1055
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 839 SDSPGSDSDSDSDSDSDSDSDSDSDSGSNNNVPPNPSPKNTGMSNKKKAAEDSNELPD 898

Qy 1056 TGANEDYGSCKTLLGTFAAGLALLGCKRRKKKK 1091
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 899 TGSEDE--ANTSLIWGLASIGSLFRKKKKKKK 932

```

RESULT 9

C)umping factor B [imported] - Staphylococcus aureus (strain N315)
F90070

C; Species: *Staphylococcus aureus*

C; date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C; Accession: F90070

R; Kuroda, M.; Ohta,

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A; Title: Whole genome sequencing of melicillin-resistant *Staphylococcus*
A; Reference number: MIRD-21311953: PMID:11418146

A;Accession: F90070

A;Status: preliminary
A;Molecule type: DNA

A;Molecule type: DNA
A;Residues: 1-877 <KUR>

A; Cross-references: C

A; Experimental source
C; Genetics:

A;Gene: clfb

Query Match

Best Local Similarity 35.6%; Pred. No. 1.9e-52;

Matches 410; Conservative 123

Qy 15

Dh

Query Match	27.28;	Score 1536;	DB 2;	Length 877;
Best Local Similarity	35.68;	Pred. NO. 1.9e-52;		
Matches 410;	Conservative 125;	Mismatches 258;	Indels 360;	Gaps 36;

Matches 410; Conservative 125; Mismatches 256

Matches 410; Conservative 125; Mismatches 258; Indels 360; Gaps 36;

```

0Y 15 IANSNKAIEKFPVAGTASVIGATLTFGLGHNDKAEKESVQDVADSNMDELSQNDQ 74
0B 8 LSNOKKISIRREFVGTSTVAGTILFGLGNHQAQASQES -----NDTQSSKNNA 59
0Y 75 SDEEKNVDINNOSINT -DDNQIIRKEETNNYDGIERKSBDRESETTNVDENATEFL 132
0B 60 SADSCKNMIEFTPO-LMTTANDTSDISANTNSANVADSTKPMSTQJNSTT---TEPAST 115
0Y 133 OKTPD---NHLBEEKKESSSVESNSSIDTFAQPSHTTINREESVQSDNVEDSHYS 189
0B 116 NETPOPTAIKQAATAAKQODTVOEANSQVD-----NNTTDANSIATNSELKNSQTL 169
0Y 190 DEANSKIKESNTSGCKEENTIEOPNKYKEDOSTSOPSGYINIDEKISMODELINPEX 249
0B 170 DLPOG-----SPQTSNMAGTKEPSVRTRARSLAVALPEPVN----- 206
0Y 250 ENKARPELSTSAQPSIKRVYVNOJLAABQGSVNHILKVTDQSTIEGYDSEGYIKAHDAE 309
0B 207 -----AADAKGTWVN-----KRVASNEKTEKTEFDEPNQSG 237
0Y 310 NLIYDVEEVDVKYSGSGDTMV-----DID-----KNVPSDLTDSFTJPKJKNDS 355
0B 238 NTFMANFTYDVKKSGSYGFTAKLPDLSITNGADVYDYSNNTMP-----IADIKSTIN 289
0Y 356 GEIATGTVDNKNKOIYTFPTYDVYDKYENIKAHILKLTYSIDKSKVPNNNTKLDVEYKTAI 415
0B 290 GGVVAKATYDILTKTYTEFVFDYVNNKENINGQESLPLFDRAKAPKSGT-YDANINIID 348
0Y 416 SSVAKTTIVEYORP-----NENKRTANLOSMFINDTKN--HTVEDTIIYNLKRKSAKETVY 469
0B 349 EMEFNKRTIYNSSPIAGIDKPGANGAISQIIGVDTAGSGNTYKOTVAFVNEPKORVILGNWY 408
0Y 470 NISGNGDEGSTIIDS-----TIKRYKVGONQLMPS-----N 503
0B 409 YIKGVODK-----IESSSKVSAVDTPKLRIFEVNTSKLSBSYVADPNDLSLKEVTQOFKN 464
0Y 504 RIYDSEYEDVTNDYDQOLGNNNDVINFNGINDSPYIIRKYSIKDIPKDKDYTT--10QYV 561

```


Best Local Similarity 27.5%; Pred. No. 5.9e-43;
Matches 461; Conservative 180; Mismatches 401; Indels 635; Gaps 60;

```

QY 7 NLLTKKP--IANKSKAIAKRFVGTASIVGATLLFGLGHNKAEENSVDYKDSNT 64
D 2 NNSKKILDFLPNKINKYISIRFVGTASIIIVGATLIFGVANDQAEAMNNTTQKODSS 61
QY 65 D-----DELSDSDSSDEKKNV INNOSINTDDNQ 97
D 62 DASKVGVNVTIEGSSANSNESDIPQVDVTKDTTEQASTEEKATTTQASTEEKADTTE 121
QY 98 IIRKEETNNVDIEK-----RSEDRTESTTNNVDENATFLQ 133
D 122 QATTEAKAEKTDKVEEAPKAEETDKATEEAPKAEIDKATEEAPKAEETDKATEE 181
QY 134 KTP---QDNHLLTEEVSESSSVSSSIDTAQPSHTTTINREESVQTSNV----- 183
D 182 EAPAEETSKATEEAPAEETSKATEEAPAEETSKATEEAPAEETSKATEEAPAEET 241
D 184 --EDSHVDFANSKIKESNTESGKEENTIEQPNKYKEDST-----TSQPS--G 227
D 242 AEETSKATEEAPKAEETNKVETEAPAEETNKATEETPAVEDTNAKSNSNAPSETE 301
QY 228 YTNIDEKISNDELINLPINEYE---NKARPLSTTS--AOPSIKRVTYNQLAEOGSNVN 282
D 302 RTQVVDVTA-KDLYKKSEVTEAKAEIEKVLPKDISNLSNEEIKKIALSEVLKANKEN 360
QY 283 HLKATYDOSTIEGVDSEGVKAKHDAENLIYDTV---FENDVKYKSGD---TMTVDIDK 335
D 361 AOPRAETFSVSANMTT-----NVNYSATLRAAODTVTKKGNTAGNDIHH 410
QY 336 NTVPDLDSFTPIKINSGEIATGT-----YDNK---NKQITYT----- 374
D 411 KTYKEEFPNEGTLTFNTFNP--NTGTGALEYNDKIDFNKDFITIVPVANNOCNTTG 468
QY 375 -----FT-----DYVDKENIKAKLTSYDKSKVPNNNTKLOVEYKFTALSUNKTI 422
D 469 ADGWGFMFTQGGQDFLNGGILR-----DKGMANASFKIDTAYNNVNGVADKL- 518
QY 423 TVEYORPRENRTANIOSM-----FTN-----IDTKNHTVEQTIYIN- 458
D 519 -----DADKTNLSQIGAAGVGYGTGVKNGADGVTVNOGNALNTMDKPNKIIYADN 571
QY 459 -----FLRYSAKETNV----- 469
D 572 TTNHLDDGFHGORLNDVLYNDATSTITATYAGTKWAKATTDGLGIDKSOKYFELITSH 631
QY 470 --NISGND-----EGSTI-----IDSTIIR-----YK 492
D 632 MONRISNIGIMRTNLEGVITTTPOADLIDDEVYTKOPIFKHKTITREDDPLLEPSPVIVOK 691
QY 493 VGDNONL-----PDSNRIDYSE-YEDVTNDYAOVL-----GNMNDVINNF- 532
D 692 GEDGKTTTPTKVPDGDVVERGEPTTEVTKNVDIEVHTPREVPQGHKDEDPMLP 751
QY 533 -----GNISPIIAYISKIDPKKDYT----- 555
D 752 IDGTEEVGKPGIKNPETGEVVT--PRVDVTKHGPKAGEBEVTKKEIPEKKKEFNPD 808
QY 556 -----TQOQVTMOTIN----- 568
D 809 LKPGGEKATQEGOGKEKTTTPTTINPLTGKVGSGEPTTEVTKPVDDELITQFGGEEVPO 868
QY 569 ---EY-----TGERTASYNNTIAFSSGOGQ----- 593
D 869 GHKDFDNLPIDTEEVGKPGIKNPETGEVVPVDDVTKHGKKAEBPEVTKKEIPE 928
QY 594 ---GLPP-----EXT-----YKGG-----DIYMEDVD----- 613
D 929 KKREPNPLKGEKEKVTQEGOTGKTTTPTTINPLTGKVGSGEPTTEVTKKEPVDDELITQ 988
QY 614 -----KDGIONTNDM-----KPLSNVLYLTITPDGSGK 642

```

```

D 989 FCGEEVPOCHKDEFPDNLPIDTEEVGKPGIKNPETGEVVPVDD--VTYHGKAGEP 1046
QY 643 SVRTDE---DGKXQPD-GIKNGLTYKITFEPEGYTPTLKHSCTNPALD---SEGN-SVM 694
D 1047 EYTKKEIPEKKEKREPNPLKPG-EKVTOEGOTGKTTTPTTINPLTGKVGSGEPTTE 1105
QY 695 VTINQODMTIDSGFYOTPKYSLGNVYDTN-----KQGIQ----- 732
D 1106 VKPEPVDIT-QFGGEEVPO---GHKDEFPDNLPIDTEEVGKPGIKNPETGEVVPV 1161
QY 733 DD-----EKGISGVKYVLKDNNGNIISTTT----- 758
D 1162 DDVTKHGKAGEBEVTKKEIPEKTRVLDPTIMEP--SPDKVAKGEBNGEKTTTPTTINP 1220
QY 759 -----DENGKY-----QFD-NLNGS----- 772
D 1221 LTGEKVGGEPTTEVTKPEIDIVNYAPELIPIHGTRIEDPLPFGERTKVIKQGLKDP 1280
QY 773 -----NYIYHFKPSGMTOTTTIDSGDDQDADAGEVHYTTIDHDFSIIDNGY 821
D 1281 ETGEIIEEPQDEVIITHGAKDSDADSDADSDADSDADSDADSDADSDADSDADSD 1340
QY 822 DDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 881
D 1341 DSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1400
QY 882 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 941
D 1401 DSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1460
QY 942 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1001
D 1461 DSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1520
QY 1002 DSDSDS-----DPSVSPSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1035
D 1521 DSDSDSDADSDSDADSDSDADSDSDADSDSDADSDSDSDSDSDSDSDSDSDSDSD 1580
QY 1036 DNDSDLGNSDSKSTKDKLPDTGANEDYSGKTLGLTFAGLAL-LLGKRRKRRNK 1091
D 1581 DADRDNHNDKTKPRNNKELPDTE--MDAQNNGTLFOSLFAALGGLFLVGRRRKKNNE 1635

RESULT 2
DSDP_HUMAN
ID DSDP_HUMAN STANDARD; PRT; 1253 AA.
AC O9NZM4; O95815;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Dentin sialoposphoprotein precursor [Contains: Dentin phosphoprotein
DE (Dentin phosphoprotein) (DPP); Dentin sialoprotein (DSP)].
GN DSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20168992; PubMed-10706475;
RA Gu K., Chang S.R., Ritchie H.H., Clarkson B.H., Rutherford R.B.;
RT "Molecular cloning of a human dentin sialoposphoprotein gene.";
RL Eur. J. Oral Sci. 108:35-42(2000).
RN [2]
RC SEQUENCE OF 463-1253 FROM N.A.
RX MEDLINE-99094526; PubMed-9879917;
RA Gu K., Chang S.R., Slaven M.S., Clarkson B.H., Rutherford R.B.,
RT Ritchie H.H.;
RL "Human dentin phosphoprotein nucleotide and amino acid sequence.";
RN [3]
RP DISEASE.

```

RX MEDLINE=21096971; PubMed=11175779;
 RA Zhang X., Zhao J., Li C., Gao S., Qiu C., Liu P., Wu G., Qiang B.,
 RA Lo W.H.Y., Shen Y.;
 RT "DSP mutation in dentinogenesis imperfecta Shields type II.";
 RL Nat. Genet. 27:151-152(2001).
 RN [4]
 RP VARIANTS DFNA39/DG11 THR-17 AND PHE-18.
 RX MEDLINE=21096982; PubMed=11175790;
 RA Xiao S., Yu C., Chou X., Yuan B., Wang Y., Bu L., Fu G., Qian M.,
 RA Yang J., Shi Y., Hu L., Han B., Wang Z., Huang W., Liu J., Chen Z.,
 RA Zhao G., Kong X.;
 RT "Dentinogenesis imperfecta 1 with or without progressive hearing loss
 is associated with distinct mutations in DSP.";
 RL Nat. Genet. 27:201-204(2001).
 CC -1- FUNCTION: DSP may be an important factor in dentinogenesis. DSP
 may bind high amount of calcium and facilitate initial
 mineralization of dentin matrix collagen as well as regulate the
 size and shape of the crystals.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed in teeth. DPP is synthesized by
 odontoblast and transiently expressed by pre-ameloblasts.
 CC -1- PRT: DSP is glycosylated.
 CC -1- DISEASE: Defects in DSP are the cause of dentinogenesis
 imperfecta 1 (DG11), also known as dentinogenesis imperfecta
 Shields type II. It is an autosomal dominant disorder in which
 both the primary and the permanent teeth are affected. It occurs
 with an incidence of 1:8000 live births. The teeth are amber and
 opalescent, the pulp chamber being obliterated by abnormal dentin.
 The enamel, although unaffected, tends to fracture, which makes
 dentin undergo rapid attrition, leading to shortening of the
 teeth.
 CC -1- DISEASE: Defects in DSP are the cause of autosomal dominant
 deafness with dentinogenesis imperfecta 1 syndrome (DFNA39/DG11
 syndrome). Affected individuals present with DG11 associated with
 early onset progressive sensorineural high-frequency hearing loss.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sb.ch/announce/>
 or send an email to license@isb-sb.ch).
 CC -----
 DR EMBL: AF163151; AAF2472.1;
 DR EMBL: AF094508; AAD16120.1;
 DR Genew; HGNC:3054; DSPP.
 DR MIM: 125485;
 DR MIM: 605594;
 JR MIM: 605594;
 KW Signal: Extracellular matrix; Glycoprotein; Sialic acid;
 KW Phosphorylation; Deafness; Disease mutation.
 KM SIGNAL 1 15
 FT CHAIN 16 1253
 FT CHAIN 16 462
 FT CHAIN 463 1253
 FT DOMAIN 439 1253
 FT SITE 488 490
 FT MOD_RES 259 259
 FT CARBOHYD 41 41
 FT CARBOHYD 49 49
 FT CARBOHYD 81 81
 FT CARBOHYD 130 130
 FT CARBOHYD 130 130
 FT CARBOHYD 150 150
 FT CARBOHYD 190 190
 FT CARBOHYD 191 191
 FT CARBOHYD 209 209
 FT CARBOHYD 222 222
 FT CARBOHYD 275 275
 FT CARBOHYD 336 336
 FT CARBOHYD 387 387
 FT VARIANT 17 17
 /FTID-VAR_012280.

FT VARIANT 18 18
 FT CONFLICT 685 690
 FT CONFLICT 744 744
 FT CONFLICT 799 799
 FT CONFLICT 836 836
 FT CONFLICT 850 850
 FT CONFLICT 868 888
 FT CONFLICT 963 963
 FT CONFLICT 1005 1005
 FT CONFLICT 1025 1025
 FT CONFLICT 1047 1047
 FT CONFLICT 1053 1053
 FT CONFLICT 1065 1065
 FT CONFLICT 1080 1080
 FT CONFLICT 1095 1095
 FT CONFLICT 1101 1101
 FT CONFLICT 1104 1104
 FT CONFLICT 1132 1132
 SQ SEQUENCE 1253 AA; 126426 MW; DAE240653904EDA CRC64;
 Query Match 12.5%; Score 708.5; DB 1; Length 1253;
 Best Local Similarity 23.2%; Pred. No. 4,4e-20;
 Matches 244; Conservative 147; Mismatches 391; Indels 269; Gaps 26;
 44 LGHNEAKAEENS-----VQDKSNTDELSDNSQDSDEKNDVYNNQSIPT 92
 436 VGHSTNGSDNSDGYDSYFDKSKMGDDPNSSDESNGNDANSESDNNSSRGDASTYS 495
 93 DDNNQIKKEETNNYD-GIEKRSDETESTTNV-----DENETFLQKTPQDNT 140
 496 DES-----KDNNGSDSKAEDDDDSDSTDTNNSDSNGNGNNDNDNDKSGDKGSDS- 549
 141 HTEFEVKRSSSVESNSSIDTAQPSHTTINRESVQSDNVEDSHVDFANSKKEEN 200
 550 --SDSSSDSSNSSDS 607
 201 TESGKEENTIEOPNKKYEESTTSOPSGYNTIDEKISNOPELLNPINEYENKARPLSTYS 260
 608 SDSSSDSDSDSDSDS-KSDSSKSE-SDSSSDSDSDSDSDSDSDSDSDSDSDSDSDS 659
 261 AOPSTIRVTYNQIAFGSGNVNHLKVTYDQSTREGYDSEGVYKAHDAENLYDVFVD 320
 660 SNS-----NSSDS-SSSS 708
 321 DKYKSGDTMTVD-IDKNYTPSDLTDSFTIPKIKDNGCELIANGTVNKKQITVFTDV 379
 709 DSSDS 768
 380 DKYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKXTALSYNKKTTVEYQRPENRRTANLQS 439
 769 DSSDS-----SDSSSDSDSDS-----NDSNSSDS 795
 440 MFTNIDTKNHTVEQTIYINPLRYSAKETVYNISGNGDEGTTIIDSTIKYKVGDNQL 499
 796 DSSNSDSDSDSDSDS-----DSSDSDSDSDSDSDSDSDSDSDSDSDSDSDS-SSNS 838
 500 PDSNRITDYSEVEDVTNDDYAOGLGNNDVNIINGNIDSPYTIKVIKYPNKKDDYTTQO 559
 839 SDSSSDS 866
 560 TYVMQTTINEYGEFTASVDNTIAFTSSGQGGGLPPEKTYKIGDYWEDVDKGIQN 619
 867 -----SDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS-SS 885
 620 TNDNEKPLSNVLTLYPPDQTSKSVRTDEGKQYDGLNGLLYKITTETPEGYPTLKH 679
 886 SDNESNS-----DSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS-DS 911
 680 SGTNPALDSEGSNVWVTINGQDDMTIDSGFYQPKYSLGNYVYDPNKGIGQDDEKGIS 739
 912 SNSDSDSDSDSDSDS-----DSSESNSSDSDSDSDSDSDSDSDSDSDSDSDSDSDS- 938
 DB


```

Db 214 GISENGTEVTPSIGDADGDDTDGSDSGNGVEDEDTGSGD----- 256
QY 257 STTSAPSIKRV---TVNOLAEQGSNVHLIVTDOSTIEGIDDS---EGVIKADAE 309
Db 257 -GEGAEAGDGRSHDGTGGCGGSHGNTDHR---GQSVSTEDDSKEOGFPNGHND 312
QY 310 NLIDVTFEVDKPKSGDTMVDIDKNTV--PSDLT-----SFTPKIKDNGELIA 360
Db 313 N-----SSENGVEEDSTQATQDKEKLSPKDRDAEGGIISGEACPSGSDQGIET 366
QY 361 TGTVDNK-NKOITFTDYDYKYENIKAKHLTSTYIDKSKYPNNNTKLDVEKTLASSVN 419
Db 367 EG--PKNKNSII-----TKESKLSGSKNSNGHQVELDKRN----- 403
QY 420 KTIIVYORPENRNTANLQSMFTNIDTKNHTVEQTIYINPLRYEAKETNVNISNGDEGS 479
Db 404 -----PKQGES-----DKPOGTAER-----SMAHNLGHSRIGSSSN 435
QY 480 TIIDSTIKYKYGDNQNLPSNRIVYSEYEDVTNDYAQLGNNNVNINFGNIDSPY 539
Db 436 SDGHS-----YEPDDESMDGDPKSDSNGSDESDT-----NSGANESGSGDASY 484
QY 540 IIRYISKYDPKDDYTTIQOTVTMOTIYNEYGEFRTASYDNTIAFSTSGGQGDLPPE 599
Db 485 TSDSSDDDDNDSGH-----AGEDD--SSDSSGGGSD----- 516
QY 600 KTYITIGDYVEDVDKGIQNT-NDNEKPLSNVLYLTYPDGTSKSVRVD--EDGKYQPDG 656
Db 517 ---SNGDSDSESEDESDSDHONS-----DSEKSDSDSDSDSDSDSDSDSDSDS 561
QY 657 LKNGITVITTEPBGYPTLKHSGTNPALSGNSVWWTINGGDMDITDSEFYQTPKYS 716
Db 562 -----SDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 604
QY 717 LGNYVWYDT-----NKDGIGQDEKIGSVKVTLKDEMN--ISTTTDENGKYOQEDN 768
Db 605 -----DTCCSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 657
QY 769 LNSGNYIHFDPKSGMOTTTDSGDDDEQADAGEVHYATTIDHDFSIDNGIYDDESDSD 828
Db 658 SDSDSD-----SSDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 709
QY 829 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
Db 710 SSDSGSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 768
QY 889 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 948
Db 769 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 828
QY 949 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1008
Db 829 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 888
QY 1009 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1051
Db 889 SKDSSSDSDSDGSKSGNGNSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 933

```

RESULT 4

```

DSDP_RAT STANDARD: PRT: 687 AA.
AC Q62598; Q9R057; P70578;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dentin staphosphoprotein precursor [contains: Dentin phosphoprotein
(Dentin staphosphoprotein) (DPP); Dentin staphosphoprotein (DSP)].
GN DSP OR R0SP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

```

```

OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM DPP-2).
RC STRAIN-Sprague-Dawley;
RA MEDLINE=20435277; PubMed=10978503;
RL Ritchie H.H., Wang L.-H.;
RT "The presence of multiple rat DSP-PP transcripts.";
RN Blochm. Biophys. Acta 1493:27-32(2000).
[2]
RP SEQUENCE OF 1-387 FROM N.A., AND SEQUENCE OF 18-26.
RC STRAIN-Sprague-Dawley; TISSUE=Odontoblast;
RA MEDLINE=94148875; PubMed=8106414;
RL Ritchie H.H., Hou H., Vels A., Butler W.T.;
RT "Cloning and sequence determination of rat dentin staphosphoprotein, a
novel dentin protein.";
RN J. Biol. Chem. 269:3698-3702(1994).
[3]
RP SEQUENCE OF 421-687 FROM N.A. (ISOFORM DPP-1).
RC STRAIN-Sprague-Dawley;
RA MEDLINE=96355551; PubMed=8702961;
RL Ritchie H.H., Wang L.-H.;
RT "Sequence determination of an extremely acidic rat dentin
phosphoprotein.";
RN J. Biol. Chem. 271:21695-21698(1996).
[4]
RP SEQUENCE OF 1-19 FROM N.A.
RC STRAIN-Sprague-Dawley;
RA MEDLINE=99333695; PubMed=10403786;
RL Yamazaki H., Kunisada T., Miyamoto A., Tagaya H., Hayashi S.-I.;
RT "Tooth-specific expression conferred by the regulatory sequences of
rat dentin staphosphoprotein gene in transgenic mice.";
RN Blochm. Biophys. Res. Commun. 260:433-440(1999).
[5]
RP SEQUENCE OF 29-33; 70-79; 93-109; 136-148; 162-188; 266-308; 398-423
AND 426-438, AND PHOSPHORYLATION OF SER-292 AND SER-298.
RC MEDLINE=21125612; PubMed=11042175;
RL Qin C., Cook R.G., Orkiszewski R.S., Butler W.T.;
RT "Identification and characterization of the carboxyl-terminal region
of rat dentin staphosphoprotein.";
RN J. Biol. Chem. 276:904-909(2001).
[6]
RP TISSUE SPECIFICITY.
RC MEDLINE=9805479; PubMed=9395101;
RA Ritchie H.H., Berry J.E., Somerman M.J., Hanks C.T., Bronckers A.L.,
Hotton D., Papagerakis P., Bernal A., Butler W.T.;
RT "Dentin staphosphoprotein (DSP) transcripts: developmentally-sustained
expression in odontoblasts and transient expression in
pre-ameloblasts.";
RN Eur. J. Oral Sci. 105:405-413(1997).
[7]
RP FUNCTION: DSP may be an important factor in dentinogenesis. DPP
may bind high amount of calcium and facilitate initial
mineralization of dentin matrix collagen as well as regulate the
size and shape of the crystals.
RC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; DPP-1/PP240 (shown here), and
DPP-2/PP171; may be produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Specifically expressed in teeth, mainly in
odontoblasts and transiently in pre-ameloblasts.
CC -1- PTM: DSP is glycosylated.
CC -1- CAUTION: Ref. 2 sequence differs from that shown due to a
frameshift in position 380.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
CC EMBL: AF247187; AAK96895.1;
CC EMBL: U02074; AAA18932.1; ALT_FRAME.
CC EMBL: U63111; AAC52774.1;
DR

```



```

FT REPEAT 513 535 23.
FT REPEAT 536 551 24.
FT REPEAT 552 573 25.
FT REPEAT 574 596 26.
FT REPEAT 597 619 27.
FT REPEAT 620 642 28.
FT REPEAT 643 665 29.
FT REPEAT 666 695 30.
SQ SEQUENCE 695 AA; 73543 MW; A670F6744CBA9B8 CRC64;

Query Match 6.9%; Score 388.5; DB 1; Length 695;
Best Local Similarity 25.0%; Pred. No. 2.6e-08;
Matches 193; Conservative 77; Mismatches 215; Indels 287; Gaps 42;

QY 469 VNISGNGDGSSTIIDSTIIKVKYK-----GDNQNLPOS---NRIYDSEYEDYNDXY 519
DB 19 LSLKAKSDGDLDDVE-INVFTDPYNAOSDGLTDLGVNKHOTHODKD-IDDD- 75
QY 520 AOLGNNDVNIINGNIDSPILIKVISKYDPNKDDYTTTIOQTVMQTTINEYGEFTASY 579
DB 76 -SIGDDEVN-NIGT-----NPKDPDSD--DGLT-DGAENVLY-----RTDPL 114
QY 580 DNTIAFSTSGGQGGDLPEPKTYKIDYWE-----DVCKDGIQNTDNKPKLSNLVTL 634
DB 115 D--ADSTTGCPMG------GAEVHRPQNGDITDDGL----- 145
QY 635 TYVDGSKSV-RTD-EDGKQFQGLKNGL---TYKITFEPEGTPTLKHSGTNPALDSE 689
DB 146 -TDGAEVNVHRTNPQDGDSDDLSDGAEVNTY-----HS--NP--KD 183
QY 690 GNSVWVTINGDQMTIDSGFYOTPKYSLGNVYVDTNKGDIQGDDEKGI-----SG 740
DB 184 GDS-----DDGVSVDGAEVNPKLK-----DSGDLTDEEELIKYRTDPCADSD 228
QY 741 VKVTLDENGNIISTTTDENGKYYQEDNLNSGNYIHFHKPSGATQTTTSGDDEQDAD 800
DB 229 FDGLLDEEYKVKHKTNPDL-DGSDDDGLDGAEVTHFN-----TNPLDASDNDGLD-D 280
QY 801 GEEVHTIITHDPEFIDNG-----YDDESDDSDSDSDSD-----SDS 839
DB 281 GEEINVHGTPEDPDSDNDGLNDEGVNVYNTPEEDSDSDGYCDGAEVNVHHTNPKDE 340
QY 840 DSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSD-----S 879
DB 341 DSDNDGIPDAEINTKHTDNDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 400
QY 880 DSDSDSDSDSD-----SDSDSDSDSDSD----- 902
DB 401 NADSDYDGLTDGAEVKRYQSNPLDKDTDDGLDGEVETVGTDPHATVTTGSRATVEI 460
QY 903 -----SD-SDSDSDSDSD-----SDSDSDSDSDSD-----SDSDS 933
DB 461 NVHGSDDNEDDTDDGLTDGAEVNLHARTDEPDATDDGLTDGAEVNTYRTNPKLASDG 520
QY 934 DSDSDSDSDSDSDSD-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD-----SDSD 982
DB 521 DGLSDGAEVNTKHSRDPDSDSDSDSDGVPDAEAKYKDSOGGLSLPTDEVREFTNPKLADTD 580
QY 983 SDSDSDSD-----SDSDSDSDSDSD-----SD-SDSDSDSDSDSD----- 1014
DB 581 FDGLTDAEAILKHKTDPRNRDTPDGDGVADGLEVTYGSDDPDADTDDGLTDGAELNVHD 640
QY 1015 -----SDSDSDSDSDSD-----SDSDSDSDSDSD-SDLGNSSDKSTKD 1051
DB 641 TNPYDADSDDDGLSDGAEVMTYHTNAKDGSDDDGKADGAEVSASTDPMWSD 692

```

```

DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Cell surface glycoprotein 1 precursor (Outer layer protein B) (S-layer
DE protein 1).
CN OLPB.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10682;
RA MEDLINE=93209931; PubMed=8458832;
RX Fujino T., Beguin P., Aubert J.-P.;
RT "Organization of a Clostridium thermocellum gene cluster encoding the
RT cellulosomal scaffolding protein C1pA and a protein possibly involved
RT in attachment of the cellulosome to the cell surface.";
RL J. Bacteriol. 175:1891-1899(1993).
CC -1- SUBUNIT: ASSEMBLED INTO MONO-LAYERED CRYSTALLINE ARRAYS.
CC -1- SUBCELLULAR LOCATION: Cell wall.
CC -1- SIMILARITY: CONTAINS 4 S-LAYER HOMOLOGY (SLH) DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X67506; CAA47841.1;
CC InterPro: IPR001119; SLH.
CC Pfam: PF00395; SLH; 3.
CC PROSITE: PS01072; SLH DOMAIN; 2.
CC Cell wall; S-layer; Signal; Repeat
CC STGNAL 1 28 POTENTIAL.
CC CHAIN 29 1664 CELL SURFACE GLYCOPROTEIN 1.
CC DOMAIN 36 763 4 X 156 AA APPROXIMATE REPEATS.
CC REPEAT 36 191 1.
CC REPEAT 207 363 2.
CC REPEAT 409 565 3.
CC REPEAT 607 763 4.
CC REPEAT 771 1377 APPROXIMATE TANDEM REPEATS OF
CC T-P-S-D-E-P.
CC DOMAIN 1378 1449 GLY/PRO/SER/THR-RICH.
CC DOMAIN 1453 1494 SLH 1 (INCOMPLETE).
CC DOMAIN 1495 1565 SLH 2.
CC DOMAIN 1566 1625 SLH 3.
CC DOMAIN 1626 1646 SLH 4 (INCOMPLETE).
CC SEQUENCE 1664 AA; 178194 MW; 5F396695BA9FE74B CRC64;

Query Match 6.5%; Score 368; DB 1; Length 1664;
Best Local Similarity 21.1%; Pred. No. 4e-07;
Matches 243; Conservative 163; Mismatches 425; Indels 320; Gaps 56;

QY 167 SHTTINRE--ESVQTSNVEDS---HVSDFANSKIKESNTE--SGKEENTIEDPNKYKD 219
DB 19 STVSNNSSFAEATPSIEMVLDKTEVHGVDTATIKNNIRKLAYOLN-----TKFD 71
QY 220 STSQPSGYTNIDKISNOBELMLPINEY---ENKARPLSTSAQPSIKRVY----- 270
DB 72 PEVLQP-----VDPATGEFTDKSPVNRLLTNSKYGK--TPVAGNDISGIITPATGY 124
QY 271 NOLAEOGSNVNH-----LIVTPOSINEGVDS---EGVKA-----HDAENT 311
DB 125 NNLTAIVSSGIDETHTGIGELGFVLLKONTSR--FEDTISMGAISGVSLEPMDAETI 182
QY 312 I-YD-----VTFEVD-DKVKSGD--TMTVID--KNVPSDLT---D 344
DB 183 TGEVIGPDLIVEAEPLKQASVALLDKTKVKYGVGDIITATIKENKNKFGVQLNIKYD 242
QY 345 SFTIPKIKDNGSELIANGTY-----DNKK-----QITVYTDVYDKENIKAH 388
DB 243 PTMLEAIELETSIAIKRTWPTGVGTVALQSDNYGKTTAVANDVAGIINFEAVYSNLFKY 302

```



```

QY 389 LKL-----TSYIDK---SKYPNNNTKIDVEKTAAL-SSVYNT-----421
Db 303 REGVAAEETGIGIGIGRRVAKAGSTAIRFEDTAMPALGEGTYMFWYGENIKGYSVQOP 362
QY 422 --IIVEYORPENNTANLQSMFTNIDTKNHTVEQTI-----YI-----457
Db 363 GEIYAESEEEGEETEPVPTETPVDPPTVTEPVPSELPSDLYVIMELDKTKKVGDDII 422
QY 458 -----NPLRYSAKETNVNISNGDGESTIIDSTIIK-VYKYGDMQNPDSNRIVDYS 509
Db 423 TATIKENMKKFAQYQINIKYDPTMLAELETGSAIAKRTWPATGGTVALQSDN-----YG 478
QY 510 EYEVYTDNDAAQAGNNNDVNNFNGNIDSPYIIKISKYDKPKDDYTIIOQTYMOTITINE 569
Db 479 KTAIVANDVGAGI-----INFAEVSN-----LTRY-----RETGVAE 511
QY 570 YTG-----EPRFASDYNT-IAFSTSGGOGDLP--EKTYYKIDGYMEDYDKGCI-----617
Db 512 ETGIIIGIKIGFVVLKAGSTAIRFEDTTPA-----MGALIEGTYMF-DWYGENIKGYSVQOP 565
QY 618 QNTNDNKRPLSNVLTLT---YPDGTSKSVRTDDEGKY-----Q 653
Db 566 EIVAEGEPEPEPVPTEPVPPTVTEPEPSELPSDLYVIMELDKTKKVGDDYIATIR 625
QY 654 FDGLKNGITYYITETPEEGTPTLKHGNTNALDSEGNVWVTING-----QDM-----703
Db 626 VNNKINLAGYQIGIK-----YDPKYLEAFNIETGQPIDEGTWPAGGTILKNRDYLPRTGVA 681
QY 704 --TIDSGFYOTPKSLGNYWYDTNKGDIQGDDEK--GISGVKVTLDENGNI-----IST 755
Db 682 INNVSKGLINLAAL-----YVFDYREGEKSEDGIIIGNIGFRY-LKAEDTIRFEELLES 736
QY 756 TTIDENGKYQFD---NLNSGNVYH-----FDKPSGMMQ-TTIDSGDD 794
Db 737 MGSIDGTYMDWMLNIRISGVYVIOPAPIKASDEPIPTDPSDEPIPSDEPTPSDEPTP 796
QY 795 DEQADAEVYVITIDHDFSIDNGYDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 846
Db 797 SDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTP 855
QY 847 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 894
Db 856 EEPITPTDPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSD 915
QY 895 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 947
Db 916 EPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPS 975
QY 948 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1000
Db 976 DEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTP 1035
QY 1001 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1054
Db 1036 SDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTP 1095
QY 1055 -----DTGANE 1060
Db 1096 EPIPTDPSDE 1106

```

```

OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9108762; PubMed=8951083;
RA Wood S.A., Ammann R.R., Brock D.A., Li L., Spann T., Gomer R.H.;
RT "RtoA links initial cell type choice to the cell cycle in
   Dictyostellium";
RL Development 122:3677-3685(1996).
CC -1- FUNCTION: MAY HAVE DUAL FUNCTIONS, ONE IN VEGETATIVE CELLS AT
   STARVATION INVOLVING CELL-TYPE CHOICE AND A LATER FUNCTION
   INVOLVING PROGRESSION TO THE TIPPED MOUND STAGE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN VEGETATIVE CELLS. LEVELS
   DECREASE DRAMATICALLY AS DEVELOPMENT BEGINS AND REMAIN LOW
   THROUGHOUT THE AGGREGATION AND LOOSE MOUND STAGES. THE LEVELS
   BECOME MORE ABUNDANT AGAIN AT THE TIGHT-MOUND STAGE AND REMAIN
   HIGH THROUGHOUT THE REST OF DEVELOPMENT AND FRUITING BODY
   FORMATION.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (see http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC CC
DR EMBL; U48298; AAC47311.1; -.
DI DictyDb; DD02056; rtoA.
KW Transmembrane; Repeat.
FT TRANSMEM 30 50
FT DOMAIN 30 38
FT DOMAIN 87 96
FT DOMAIN 181 181
FT REPEAT 181 190
FT REPEAT 191 200
FT REPEAT 201 211
FT REPEAT 212 222
FT REPEAT 223 233
FT REPEAT 234 244
FT REPEAT 245 255
FT REPEAT 256 266
FT REPEAT 267 277
FT REPEAT 278 287
FT REPEAT 288 298
SQ SEQUENCE 400 AA; 39840 MW; 0392FE4E68DC27A75 CRC64;
Query Match 6.4%; Score 361.5; DB 1; Length 400;
Best Local Similarity 30.1%; Pred. No. 1.4e-07;
Matches 100; Conservative 48; Mismatches 123; Indels 61; Gaps 5;
QY 744 TLNDENGNITSTYTTDENGKYQFNDLNSGNVYHFDKPSGMMQTTTIDSGDDDEQADAGEE 803
Db 49 TIKKKNSSISS-----SQGLSEVDSIDISSG-- 77
QY 804 VHYTTIDHDFSIDNGYDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 863
Db 78 -----SNSTYASSEGSYSSSSSSNGSGOSTSNGSGEASGSSNGSGOSTSNGSGEFA 124
QY 864 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 923
Db 125 SGSSNSGSGOSTSDSSNGSGOGSTGSSNGSGOSTSDS-SNGSGOGSTSDSSNGSGOGSTGSS 183
QY 924 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 983
Db 184 NSGSESSGSSNSGSESSGSSNSGSESSGSS-SNGSGSESSGSSNSGSESSGSSNSGSESS 242
QY 984 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSLGN 1043
Db 243 SSGSSNSGSESSGSSNSGSESSGSSNSGSESSGSSNSGSESSGSSNSGSESSGSSSDSGS 302
QY 1044 SSDKST-----KDKLPDTGANE 1060

```

DB 303 SSDGKTTCTISFHDLTSLINTVDDDEICTGKE 334

RESULT 11
Y67 YEAST
ID Y67 YEAST STANDARD: PRT: 1658 AA.
AC 003661: 004988:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 Intergenic region.
GN YMR219W OR YMR261.13 OR YMR959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
ON NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-711 FROM N.A.
RC STRAIN=S288c / AB972;
RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 608-1648 FROM N.A.
RC STRAIN=S288c / AB972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: 249809; CA98934.1; -
DR EMBL: 249933; CA90190.1; -
DR SGD: S0004832; YMR219W.
KM Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3693F968305A757D CRC64;

Query Match 6.18; Score 343.5; DB 1; Length 1658;
Best Local Similarity 19.2%; Pred. No. 3.3e-06;
Matches 241; Conservative 215; Mismatches 446; Indels 355; Gaps 57;

QY 47 NEAKAEENSVDYKSDTDELSDSNDOSDDE-----KKNVINNNNSINDDNNQI 101
DB 232 DEEYAEAGALQDY-----SNDYAELEGQVERKN--IGQEQN-NVENATQISS 277
QY 102 E--ETNNY-DGIEKRSDEPTSTTNVDENATFLQKTPQ-----DNT----- 140
DB 278 DSEGGQVSGVEEMLEDDIDVESDAKDESGAEGTEHSVD?SKYQPRDNKIVIE 337
QY 141 -----HLTEEVKESSVE--SSNSIIDTAQOPSHTTI-----NREESVOTS 180
DB 338 KYESDEKRVHQRYSEDGAFGFSVNIISVDDSEDEESQAESYSANAENVYHHNHELDK 397
QY 181 DNVEDSHVDFPANKIKESNTESKEENTIEQPKVKEDSTTSQPSGT--NIDEKISNQ 238
DB 398 ELLEDISSSESOSAES--EAGSE--DFFYKMKNEKSTSEETESRQGR-K 451
QY 239 DELLNLPINEYENKARPLSTTSAPSI-KRYTVNQLAAEQSNVNLIKYVDOSITE-- 294
DB 452 DAYTKNVVEQENDEPEKDOIIRSSIDKFNHNNKSEYENV--LENEIDPAIVEREN 509
QY 295 GYDSDSEGVIAHADEANLIVDTFEVDKVKSGTMTYDIDKNYPSDLTDSFTTPKIDN 354
DB 510 QINDVEG-----YDVT-----GKSVESDLHSHS--PDNLXYDAARALQFQ 548
QY 355 SGEIIATGTDNKNKOITVPTDYVD-----KYENIKAKILKLSYDKSKVNNNTKL 407
DB 549 QSR---NSNCPQKEEQVSESTLGHNSGNSNLGRLDSESEQIPLKDTTGEN--NNNLKT 602

QY 408 DVEKTAALSSVKNITTYEYORPENRPAANQSMFTNIDTAKNHYEOTIYINPLRYAKET 467
DB 603 D-----RGDLSSSVLEIEVKVSEKRLDG-----STERELY-----PL---STD 638
QY 468 NVNISGNGDGSRT--RDDSTIIKVVYVGDNONLPD-----SRIIDYS 509
DB 639 TINNSLGNEDSIYSLDDADAL-----SENLTDPPLMEIKTTPREYVYSESYSST 691
QY 510 EYEDVT-----NDYVLAQGNNDYNIN-----FGNIDSPYI 540
DB 692 SYEDNTYAMPQVEYTPSPFNNDPNSLNDDEYK---KHDLKSTLALAPAFKKDAFEV 748
QY 541 IKYISK-----YPPNKD-----DYTTIQOTVMTQITNEYTGEFRASIDNTI 583
DB 749 EAGVTAKSLTSTSGHTNIFHTSKETKQVSDLDESTENVTFE---NENYGDENKQSKNFP 805
QY 584 AFTSSGQGGDLPPEKTYIAGDVWEDVAKDQGNQNDNEKPLSNLYVLTLPD---GT 640
DB 806 GVANSTDKSTEDNTDEYFSAINr---TWTGDSCEDIETASVNEENLRCEKDMNE 861
QY 641 SKSVRTDEDKYQFDGLKNGLTLYKITPEPEGYPTLKHGQTNPALDSEGSVWVTINGQ 700
DB 862 AEMSGDECYKQNDGSKTQISF--STDSPDNr---QESNDN---TESSSTYKVRNS 911
QY 701 DDMTIDSGFYQTPKYSLGNYVYDTNKDIO--GDEKGISGVYTLKDE--NGNIIST 755
DB 912 DLEDDELKKEITLKAELVAVVLDKEESESDEYQVADDEPG-----NDEGSNENIVKG 963
QY 756 TTIDENKRYQFDN-----LNSGNYIYHPKPSGMQTTIDSDDD 795
DB 964 TKKDTLIVPEENKRVKVEEETLFEANVSSVYONKMDHTDVINDQANVAGEERK 1023
QY 796 E--ODADGEEVHYITDh--DDFSIDN----- 818
DB 1024 YTIQNTPEKHIITIERIDENALGNMEIPERCVCETKTHNEVLFERRATIENTKALEN 1083
QY 819 --GYDDESDDS 873
DB 1084 NTNMDQVQACSDS 1143
QY 874 DS-----SDS 903
DB 1144 SSPIRIGAVVYGVKYVDAEVSFKRIDVNDSDSDNDVIGDYNQDIFKNSNS--TDASV 1202
QY 904 DS 949
DB 1203 NMAVSSKERSDDEDAVILGVTAEAHANDGNNSRVINIDPTNGAYEEDSEVFRQOYK 1262
QY 950 DS 1009
DB 1263 DKENLHKSEEPVLEGLQSEQHFEKKDHSENEEDDTIYGOITSANIHSHNAPDDIKROQL 1322
QY 1010 DSVSDSDSDS-----DSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1061
DB 1323 KNLSDLENYSQLIEDSRGKNOESDESVNRSRDL--TFEKSVMKEY--AGAIEE 1375

RESULT 12
V648_HSVSA
ID V648_HSVSA STANDARD: PRT: 797 AA.
AC 001033;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Hypothetical gene 48 protein.
GN 48 OR EDLF5.
OS Herpesvirus salmatri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
ON NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.

RESULT 13	
ALAI_CANAL	
ID	STANDARD;
ALAI_CANAL	PRT; 1419 AA.

DB 211 -----SGTMCSTSYGDV-AIDCSNVHIGISKGVANDMNNHPPTSSESYTKSCSPGISI 263

OY 334 DKNTVPS--DLTDSFTIPKIKDNGSEIIATGTYYDNKNKKQITYYTFDYV---DKYENIKA 387


```

DB 264 TQONNPAGRRPFDATISPS-----DNNQIOLSK-NDYTCVDDYKOHAPP 308
OY 388 HLLTSYIDKSKVYNN-----MTKLDEYKALS-----SVNKITVEYORPENRNTAN 436
DB 309 TLKMGYKNSDAGNSGIVATRTYTDSTAVTTLPPNRSVAKTIELQIP---TTT 366
OY 437 LOSMETNIDTKNHTVEQTIYINPLRSKAKETNNVNSGNGDEGSIITIDSTIIVYKVDN 436
DB 367 ITTSYGVVTS-----YSTKTAPI-----GETATVIYD----- 394
OY 497 QNPDSNRIDYSEYEDVNDYAOGLNNNDVAINNGINDSPYIIVYISKYDKNKDYT 556
DB 395 -----VPHYTTTTVSEWGTITTTTTTRNPTDSIDT-VVOQPS---PNPTTTT 441
OY 557 --IQOTVMTQTINE-----YTGEFRTASYDNTIATFSTSGGCGDL 596
DB 442 QPMSESEFTSTTTITNLSKGDSDIVREPHNPVTTEFEFSEFATFETISK----- 493
OY 597 PPEKTKIDYWEVDYDKGDIQNTNDEKPLSNVLV-----TLTPDGTSKSVRDEDEG 650
DB 494 -PEGT-----DSVIVREPHNPVTTEFEFSEFATFET- 525
OY 651 KYGFDELKNGLYKITEFPEGTPTLKHSGTNPALD-----SEGNSVWTI---NGOD 701
DB 526 -----ITNG-----PEGTSDVIVREPHNPVTTEFEFSEFATFETITNKPEGTD 570
OY 702 DMTIDSGF---YOTPKYSLGNVYVDTNKGDIQDEKIGSKVYTLKDENGNIISTT- 756
DB 571 SVYKEBYNPTVTTTEFMSVATFETITNGPEGTD-----SVIVREPHNPVTTEF 623
OY 757 -----TTDENGKYQFDNLNGNIVHEDKPSGKTOTTTSCDDDECDADGEVHVTITD 810
DB 624 WSESYATTE-----TITGPGKTD-----SIVI 646
OY 811 HDPEFSDNGYDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 870
DB 647 HDPLESSSTTALES-SDSNISSAOGSSSVESSSIVGLSSSDIPLSSDMPSSTG 705
OY 871 -SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 926
DB 706 LTTSESSTVSYSDSDSSSTLSSSDRCSSISDPTTFPMD-SSSDLESTITWSSS 764
OY 927 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 986
DB 765 IDAQSHLVQSVNSISTGSELSSSSSEESTFATDVLVSADASSILSPSSVYPSSTI 824
OY 987 SDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1038
DB 825 SSSDDPHTIAGEBDSLSISFTISTVEISDSVSLTDPASSFSSSSSLNDSSSSPSSD 884
OY 1039 -SDLGNSDKST 1049
DB 885 QSDILTSSSFST 896

```

```

RESULT 14
HYRL CANAL
ID HYRL CANAL STANDARD; PRT; 937 AA.
AC P46591;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hyphally regulated protein precursor.
GN HKR1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_Taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RX MEDLINE=96404781; PubMed=8808922;

```

```

RA Bailey D.A., Feldmann P.J.F., Bovey M., Gow N.A.R., Brown A.J.P.:
RT "The Candida albicans HYRL gene, which is activated in response to
RT hyphal development, belongs to a gene family encoding yeast cell wall
RT proteins."
RL J. Bacteriol. 178:5353-5360(1996).
CC -1- FUNCTION: NONESSENTIAL COMPONENT OF THE HYPHAL CELL WALL
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- DEVELOPMENTAL STAGE: ABUNDANT IN HYPHAE.
CC -1- INDUCTION: INDUCED SPECIFICALLY IN RESPONSE TO HYPHAL DEVELOPMENT.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; Z50123; CA90485.1;
KW Cell wall; Glycoprotein; Signal; GPI-anchor; Transmembrane; Repeat.
FT SIGNAL 1..20
FT CHAIN 1..913
FT PROPEP 914..937
FT TRANSMEM 42..62
FT DOMAIN 332..576
FT DOMAIN 569..576
FT DOMAIN 577..814
FT DOMAIN 610..753
FT REPEAT 610..613
FT REPEAT 666..669
FT REPEAT 680..683
FT REPEAT 690..693
FT REPEAT 698..701
FT REPEAT 738..741
FT REPEAT 750..753
FT DOMAIN 901..904
FT LIPID 913
FT CARBOHYD 16..16
FT CARBOHYD 236..236
FT CARBOHYD 449..449
FT CARBOHYD 488..488
FT CARBOHYD 580..580
FT CARBOHYD 585..585
FT CARBOHYD 595..595
FT CARBOHYD 603..603
FT CARBOHYD 619..619
FT CARBOHYD 631..631
FT CARBOHYD 641..641
FT CARBOHYD 649..649
FT CARBOHYD 711..711
FT CARBOHYD 747..747
FT CARBOHYD 759..759
FT CARBOHYD 773..773
FT CARBOHYD 897..897
FT CARBOHYD 913..913
SQ SEQUENCE 937 AA; 93699 MW; 17802F121E5BA926 CRC64;

Query Match 5.64; Score 315; DB 1; Length 937;
Best Local Similarity 21.58; Pred. No. 2,1e-05;
Matches 189; Conservative 133; Mismatches 349; Indels 208; Gaps 30;

```

```

Db 149 -----LIVAVO--NOKAGNIAR-----GTAVOTITNNQICLRHOF 184
QY 466 ETNNISNG-----DE-----GSTII-----DSTIIKYVVGNO----- 497
Db 185 VPAKIKGTGVTADDEDTMIRKGLTILVEPTNHFYLDKSSSLIVHAVSSNOTFTVGF 244
QY 498 -----NLP-DSNRIYSEYEDVTN-----DDYQALNNNNVNFNIDPSY 539
Db 245 GNGKGLITLPLTGNRHRHFFREYTPDGLDRLDALPOLYKIKGIDSKL-FRIVSRG 303
QY 540 IIVYISKYD---PKKDYTTIQOTVMTQITINEYTGFEFTASNDYNTAFSTSSGOGGL 596
Db 304 LKNAYT-YDGVVNNNEIPAVCLICTNPSAPESESDLNTFTTSIETSSVSSATEBSV 362
QY 597 PPEYTIKGVWEDVDKDGQNTN-DNEKPLSNVLTLPDGTSAKVRTEDECKIOFD 655
Db 363 VSESSAAV-----DSLTSLSLSKSSSVSVSTTIESSSTAFETNNSESTD 412
QY 656 GLKNGLTAFETPEGYPTLKSGTNPALDSEGSVWVWVINGODMTIDSGFYQPKY 715
, 413 AGSSIS-----QSESSSALTSSSETS---SSESMASSTASNTSIETSGIYSQS-- 462
QY 716 SLGNVYDINKDGIQGDDEKISGVKYLKDENGNIISTTTDENGKYQFDNLN----- 770
Db 463 -----ESSSNALSTEQSITSSPGSTIVNSTVSTITSDENKCTEDVYTIFTV 514
QY 771 -----SGNY-----IVHFKPS-----GMTOTTT 789
Db 515 PCSTDCVPTTGDIEMSTSYRTORTVSTFTTNCDEVSCQDVYTYTNNVPHTVDAITTTT 574
QY 790 DSGDDEDDAGEVHYHTIDHDFSIDNGYYDESD-SDSDSDSDSDSDSDSDSDSDSD 848
Db 575 STGDNS-----TGNGSGSNHGGNGSTEGSGSGSAGSNEGSGSPNNNSG 622
QY 849 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 908
Db 623 SGSEGGSNNGSGSDSGSGSNNGSGSGSTEGSEGGSGSGSDSGSGSDSGSGSDSG 682
QY 909 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 968
Db 683 SEGSGSGNEGSGHSGNEGSGSGSGSNNGSGSGSQSGSGSGSGSGSGSGSGSGSN 742
QY 969 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1028
Db 743 PGAGSGNEGSGGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 802
QY 1029 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1061
, 803 SGSKAGSHSGSGNEGAKTDSIEGFHTESK-PGFNTGAHTD 840

```

```

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
CC TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST
CC COMPONENT OF RNA POLYMERASE II.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: THE TANDEN 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16561; CAA34560.1; -
DR PIR: S07485; RNZ02L.
DR InterPro: IPR000684; RNA_polII_repeat.
DR InterPro: IPR000722; RNA_pol_A.
DR InterPro: IPR002879; RNA_pol_A2.
DR Pfam: PF00623; RNA_pol_A; 1.
DR Pfam: PF01854; RNA_pol_A2; 1.
DR PROSITE: PS00115; RNA_POL_II_REPEAT; 9.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN-FING 68 411
FT DNA_BIND 378 725
FT DOMAIN 707 1128
FT DOMAIN 1093 1128
FT DOMAIN 1144 1159
FT DNA_BIND 1182 1193
FT DOMAIN 1687 1694
FT DOMAIN 1258 1290
FT DOMAIN 1261 1290
FT DOMAIN 1602 1612
FT DOMAIN 1746 1759
FT DOMAIN 1806 1820
FT DOMAIN 2061 2246
FT DOMAIN 2247 2384
FT DOMAIN 2452 AA; 278166 MW; F995E117F617A48F CRC64;
SEQUENCE
Query Match 5.58; Score 311; DB 1; Length 2452;
Best Local Similarity 19.18; Pred. No. 8.6e-05;
Matches 247; Conservative 211; Mismatches 458; Indels 374; Gaps 55;
QY 54 NSVDVKSNDDELS-----DSNQSDEEKNDVYN 85
Db 659 NSTND-----NTDDYLEKKNAYSRSNGNHNPSLSTGDIINNGVNOQNMSSPNNNN---N 711
QY 86 NNSGINTDDNNQIIKKEFTNNDG-----IEKRSDEPTESTTNDVNEAFFLQ--K 134
Db 712 NNNNNNNNNNN-----NNNIGGINSFKFRFNMYKINLMRDSSTPSKDDNPYCSINDGK 764
QY 135 TPQDNTHTLEEVKSSSVSSNSI-----DPAQDSHTTINREESVQISD 181
Db 765 VIKNNELSLGIIKRTYSGSSGLIHVLMHEMGDPKTKFLSALQKVTNNMLEVGTIV 824
QY 182 NVEDSHVDFANSKIKESNTESGKEENTIEOPNKVKEDSTSQP--SGYTNIDEKISNOD 239
Db 825 SCSDIASNKVLGAVRELIDSKSESVKIVE--KQKSELCEQPKSKLSYSEFETRVNNE- 861
QY 240 ELLNPLINEYENKARPLSTSAQPSI-KRYTVNOLAA--EGGS--NVNHLIKYVDOSTE 294
Db 882 -----LNCAREMAGVASESIDERNINIFSVAGSGSIINISQIISVCVQOQWVE 931
QY 295 G-----YD-----DSEGVIRKHAENL-LYDTPE-----VDDKYS 325

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 41.7491 Seconds

(without alignments)
5389,420 Million cell updates/sec

Title: US-09-147-405b-15

Perfect score: 5646
Sequence: 1 MINKNNLTKKKPIANKSN.....FAGLGALLGKRRNRKNKN 1092

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

SPTREMBL.21.*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-rylirus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2	070022
2	4485	79.4	931	2	09K113
3	2756	48.8	1171	2	09KWX6
4	2698.5	47.8	1166	2	086489
5	2589	45.9	1141	16	099W46
6	2586	45.8	1141	16	0932F7
7	2303.5	40.8	1733	2	09K114
8	2234	39.6	1385	16	099W47
9	2093.5	37.1	953	16	099W48
10	2023.5	35.8	947	2	086487
11	1880	33.3	1315	2	086488
12	1652	29.3	989	16	099VJ4
13	1628	28.8	933	2	053653
14	1610	28.5	913	2	086476
15	1581	28.0	935	16	0932C5
16	1356	27.2	877	16	099R07

17	1471	26.1	881	2	093MH7
18	1216	21.5	1698	2	09LC00
19	785	13.9	882	16	092AK9
20	782.5	13.9	903	16	08Y697
21	671.5	11.9	487	2	09K112
22	670.5	11.9	970	11	08VBY1
23	624.5	11.1	2271	16	099QY4
24	612	10.8	2283	2	08VQ99
25	554	9.8	3394	5	077384
26	549	9.7	2570	2	092FE9
27	547	9.7	2276	2	093YV6
28	525.5	9.3	1038	16	099RD2
29	520	9.2	940	2	053682
30	502.5	8.9	961	16	099RD3
31	499	8.8	840	5	095YF6
32	498	8.8	451	3	09P785
33	471	8.3	334	5	077406
34	458.5	8.1	1946	5	097291
35	452	8.0	2910	5	026008
36	450.5	8.0	1514	5	085Y53
37	449	8.0	521	17	08T545
38	443.5	7.9	1301	5	08WSK5
39	442.5	7.8	1999	16	08YWB9
40	442	7.8	691	4	096OF7
41	441.5	7.8	1114	5	097242
42	441	7.8	2402	2	09AER7
43	439	7.8	1461	5	095259
44	435	7.7	1192	5	096127
45	430.5	7.6	4550	5	077336

ALIGNMENTS

RESULT 1	070022	PRELIMINARY;	PRT;	1092 AA.
ID	070022			
AC	070022;			
DT	01-AUG-1998 (TREMBLrel. 07, Created)			
DI	01-AUG-1998 (TREMBLrel. 07, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Fibrinogen-binding protein precursor.			
OS	Staphylococcus epidermidis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;			
OX	NCBI_TaxID=1282;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HB;			
RX	MEDLINE=98261511; PubMed=9596732;			
RA	Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;			
RT	"A Fibrinogen-binding protein of Staphylococcus epidermidis."			
RL	Infect. Immun. 66:2666-2673(1998).			
DR	EMBL: Y17116; CAA76638.1; -			
DR	InterPro: IPR001899; Gram_pos_anchor.			
DR	Pfam: PF00746; Gram_pos_anchor; 1.			
DR	TIGRFAMS: TIGR01167; LPRG_anchor; 1.			
DR	TIGRFAMS: TIGR01168; YSIRK_signal; 1.			
DR	PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.			
KW	Signal.			
FT	SIGNAL 1 51			POTENTIAL.
FT	CHAIN 52 1092			FIBRINOGEN-BINDING PROTEIN.
FT	SEQUENCE 1092 AA; 119292 MW; 6542BC39AAB8B984 CRC64;			
QY	Query Match	100.0%;	Score 5646;	DB 2; Length 1092;
	Best Local Similarity	100.0%;	Pred. No. 1.5e-182;	
	Matches 1092;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1 MINKNNLTKKKPIANKSKVAIRKFTYGTASIVIGATLLFGIGHNEAKAEENSVDYK 60			
DB	1 MINKNNLTKKKPIANKSKVAIRKFTYGTASIVIGATLLFGIGHNEAKAEENSVDYK 60			
QY	61 DSNFTDELSDSDSSDEEKNDYVYNNQSGINTDNNQIITKEETNNIDGIEKRSERTES 120			

```
Db 61 DSNDDLSDDSDSDSEKNDVINNOSINTDDNOIIRKKEETNNVIGIKRSDRES 120
Oy 121 TTVNDENATFLQTPDQNTHLTEEVKSSVSSNSIDTAOOPSHTTINRESVOTS 180
Db 121 TTVNDENATFLQTPDQNTHLTEEVKSSVSSNSIDTAOOPSHTTINRESVOTS 180
Oy 181 DNVDSHVSDFANSKIKESNTESGKEENTIEOPKVKREDSTTSQPSGTINDEKISNDE 240
Db 181 DNVDSHVSDFANSKIKESNTESGKEENTIEOPKVKREDSTTSQPSGTINDEKISNDE 240
Oy 241 LNPINEYENKARPLSTTSQPSIKRYVNOQLAEOGSNNHLIKYTDOSITEGYDSE 300
Db 241 LNPINEYENKARPLSTTSQPSIKRYVNOQLAEOGSNNHLIKYTDOSITEGYDSE 300
Oy 301 GVIAHAENLIYVTEFVDKVKSGDTMVIDKNTVPSDLTDSFTIPKIKDSGELIA 360
Db 301 GVIAHAENLIYVTEFVDKVKSGDTMVIDKNTVPSDLTDSFTIPKIKDSGELIA 360
Oy 361 TGTVDNKKOITTYFTDVKYENIKALKITSYIDSKVYNNNTKLDVEYKTAALSSVKN 420
Db 361 TGTVDNKKOITTYFTDVKYENIKALKITSYIDSKVYNNNTKLDVEYKTAALSSVKN 420
Oy 421 TTYEYORPENRTANLOSMTNIDTKMTEQTIYINPLRYSAKETNVNISGNGDEST 480
Db 421 TTYEYORPENRTANLOSMTNIDTKMTEQTIYINPLRYSAKETNVNISGNGDEST 480
Oy 481 IIDDSTIIKVKVGDNONLPDSNRIDYSEYEDVTNDYAOAGNNDVNIENFGNIDSPYI 540
Db 481 IIDDSTIIKVKVGDNONLPDSNRIDYSEYEDVTNDYAOAGNNDVNIENFGNIDSPYI 540
Oy 541 IKVISKIDPKNDYTTIOQTVMQTTINEYGEFRASVDTIAPSTSSGGGQDLPPEK 600
Db 541 IKVISKIDPKNDYTTIOQTVMQTTINEYGEFRASVDTIAPSTSSGGGQDLPPEK 600
Oy 601 TYKIGDYVMEVDKGIQNTDNDNEKPLSNVLTLPDGTSGSVRTDDEKGYOGDGLKNG 660
Db 601 TYKIGDYVMEVDKGIQNTDNDNEKPLSNVLTLPDGTSGSVRTDDEKGYOGDGLKNG 660
Oy 661 LTYKITEEPBGYTPPLKHSCTNPALDSEGSVWVTINGQDMDTIDSGFYOTPKYSLGNY 720
Db 661 LTYKITEEPBGYTPPLKHSCTNPALDSEGSVWVTINGQDMDTIDSGFYOTPKYSLGNY 720
Oy 721 VWYOTNKGIGQDEKIGSGYKVTLKDENGNIISTTTTDENGKYQFNLNGNYYVHFDK 780
Db 721 VWYOTNKGIGQDEKIGSGYKVTLKDENGNIISTTTTDENGKYQFNLNGNYYVHFDK 780
Oy 781 PSGMTQTTTDSGDEQDADGEVHYTTIDHDFSIDNGYTDDESDSDSDSDSDSDSDSD 840
Db 781 PSGMTQTTTDSGDEQDADGEVHYTTIDHDFSIDNGYTDDESDSDSDSDSDSDSDSD 840
Oy 841 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 900
Db 841 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 900
Oy 901 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 960
Db 901 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 960
Oy 961 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 1020
Db 961 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 1020
Oy 1021 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 1080
Db 1021 SDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDS 1080
Oy 1081 LGGRRKNKNKN 1092
Db 1081 LGGRRKNKNKN 1092
```

RESULT 2

```
09K113
ID 09K113 PRELIMINARY: PRT: 931 AA.
AC 09K113;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fibrinogen-binding protein stdg.
CN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID:1282;
RX STRAIN-K28;
RX MEDLINE-20340957; PubMed-10878118;
RA McCreia K.W., Hartford O., Davis S., Ni Eidin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Str) protein family in Staphylococcus
epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245042; AAF72510.1; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor.
DR TIGRFAMS: TIGR01167; LPTXG_anchor.
DR TIGRFAMS: TIGR01168; YSIRK_signal.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SO SEQUENCE 931 AA; 102955 MW; 591657D97027116 CRC64;

Query Match 79.4%; Score 4485; DB 2; Length 931;
Best Local Similarity 81.1%; Pred. No. 1,3e-143;
Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

Oy 4 KNNLLTKRRKRIKANKSNKYAIRKFTVGASIVIGATLLFGHGHEAKAEKNSVOVDKSN 63
Db 3 KNNLLTKRRKRIKANKSNKYAIRKFTVGASIVIGATLLFGHGHEAKAEKNSVOVDKSN 62
Oy 64 TDDELSDSDSDSDSEKNDVINNOSINTDDNOIIRKKEETNNVIGIKRSDRESSTTN 123
Db 63 MDDLSDDSDSDSDSEKNDVINNOSINTDDNOIIRKKEETNNVIGIKRSDRESSTTN 121
Oy 124 VDENEATFLQTPDQNTHLTEEVKSSVSSNSIDTAOOPSHTTINRESVOTSNNV 183
Db 122 VDENEATFLQTPDQNTHLTEEVKSSVSSNSIDTAOOPSHTTINRESVOTSNNV 181
Oy 184 EDHSVDSFANSKIKESNTESGKEENTIEOPKVKREDSTTSQPSGTINDEKISNDELLN 243
Db 182 ENSVDSFANSKIKESNTESGKEENTIEOPKVKREDSTTSQPSGTINDEKISNDELLN 241
Oy 244 LPINEYENKARPLSTTSQPSIKRYVNOQLAEOGSNNHLIKYTDOSITEGYDSEGYI 303
Db 242 LPINEYENKARPLSTTSQPSIKRYVNOQLAEOGSNNHLIKYTDOSITEGYDSEGYI 301
Oy 304 KAHAENLIYVTEFVDKVKSGDTMVIDKNTVPSDLTDSFTIPKIKDSGELIATGT 363
Db 302 KAHAENLIYVTEFVDKVKSGDTMVIDKNTVPSDLTDSFTIPKIKDSGELIATGT 361
Oy 364 YDNKNOITTYFTDVKYENIKALKITSYIDSKVYNNNTKLDVEYKTAALSSVKNYIT 423
Db 362 YDNKNOITTYFTDVKYENIKALKITSYIDSKVYNNNTKLDVEYKTAALSSVKNYIT 421
Oy 424 VEYORPENRTANLOSMTNIDTKMTEQTIYINPLRYSAKETNVNISGNGDEGSTIID 483
Db 422 VEYORPENRTANLOSMTNIDTKMTEQTIYINPLRYSAKETNVNISGNGDEGSTIID 481
Oy 484 DSTIIKVKVGDNONLPDSNRIDYSEYEDVTNDYAOAGNNDVNIENFGNIDSPYIITV 543
Db 482 DSTIIKVKVGDNONLPDSNRIDYSEYEDVTNDYAOAGNNDVNIENFGNIDSPYIITV 541
Oy 544 ISKYDPKNDYTTIOQTVMQTTINEYGEFRASVDTIAPSTSSGGGQDLPPEKTYK 603
Db 542 ISKYDPKNDYTTIOQTVMQTTINEYGEFRASVDTIAPSTSSGGGQDLPPEKTYK 601
```

QY	IGD	YVME	VDKDG	IGONT	NDNEK	PLSNV	VTLT	YTPD	GTSTK	SVRTD	DGKYQ	PDGL	KNG	LTY	663						
604	IGD	YVME	VDKDG	IGONT	NDNEK	PLSNV	VTLT	YTPD	GTSTK	SVRTD	DGKYQ	PDGL	KNG	LTY	663						
602	IGD	YVME	VDKDG	IGONT	NDNEK	PLSNV	VTLT	YTPD	GTSTK	SVRTD	DGKYQ	PDGL	KNG	LTY	661						
664	KIT	FET	PEG	YPT	LKHS	GTN	PALD	SEGN	SVW	YTT	ING	ODM	ITD	SGFY	TPPY	SG	YV	723			
662	KIT	FET	PEG	YPT	LKHS	GTN	PALD	SEGN	SVW	YTT	ING	ODM	ITD	SGFY	TPPY	SG	YV	721			
724	DTN	KDG	IG	ODDE	KEG	IS	GVK	YTL	KD	ENG	NI	IST	TTT	DENG	KYQ	FD	LN	SG	YV	783	
722	DTN	KDG	IG	ODDE	KEG	IS	GVK	YTL	KD	ENG	NI	IST	TTT	DENG	KYQ	FD	LN	SG	YV	781	
784	MTQ	TTT	SG	DD	ED	DAD	EE	VH	YTT	HD	DS	IND	NGY	YD	DD	SD	SD	SD	SD	SD	843
782	MTQ	TTT	SG	DD	ED	DAD	EE	VH	YTT	HD	DS	IND	NGY	YD	DD	SD	SD	SD	SD	SD	834
844	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	903
835	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	834
904	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	963
835	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	834
964	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	1023
835	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	881
835	DS	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	881
1024	GSD	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	1083
882	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	922	
1084	RR	NR	KN	KN	1092																
923	RR	NR	KN	KN	931																
RESULT 3																					
09KMX6																					
ID	09KMX6																				
AC	09KMX6																				
DT	01-OCT-2000																				
DT	01-OCT-2000																				
DT	01-JUN-2002																				
DE	Bone																				
GN	BAP																				
OS	Staphylococcus aureus																				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales																				
XC	Staphylococcus																				
JX	NCBI_TaxID=1280;																				
RN	[1]																				
RP	SEQUENCE FROM N.A.																				
RC	STRAIN=024;			</																	

[illegible]

[illegible]

[illegible][illegible]


```

QY 179 TSDNVEDSHVDFANSKIKESNTESGKEENTIEQPN-KVKEDSTTSQPSGYTNIDEKISN 237
D 158 -PDLEENKSVVN-----VQPTNEENKRVDAKTESFTLVNKSDAIKS-----NAETLVDN 205
QY 238 ODELLNLPINEYENKARLSTTSAPPSIKRVTVNOIAEQQS-----NNHLIKYTDOST 293
D 206 NSNSNN-----ENNADILPKSTAPKSLNTRMAAIOQPNSTDSKNVNDLSTNTTLTV 259
QY 294 EGYDSEGVAKAHOENLIYDTPFEVDKVKSGDMTMDIDKNTY-----PSDLTDSFT 347
D 260 VDAHQNSKTIYPAODYLSKSOIT--VDDKVKSGDFTIKY--SDTYOYVGLNPEDKN-- 313
QY 348 IPKIKD-NSCEIATGYDNKNKOITTYTDYVCKENIKAKLKTYSIDSKVYNNNTK 406
D 314 IGDIKDPNNGETIAFAKADTANNLITYFTDYVDFNSVKKGLNLSITMDATIPVD--K 371
QY 407 LDVEYKVALSSVKNITTYEYORP--NENRTANLOSMTNIDTKNHT-----VEQTIY 456
D 372 KDVEFSTVIGQITTTTADITYPAYKEADNNSIGSAFT--ETVSHGVNVEDPGYVQYVY 429
QY 457 INPLRSKKEINNVNNSGDEGST-----IIDSTIIKYYKKGDNONLDSNRIDY--SE 510
D 430 VNPMDKLGAKLKEAYHPKPTNIGQINONVNIKIRVPEGYTL--KKGVDVNTND 486
QY 511 YEDVTND--DYAOLGNNDVNIENFGNIDSPYIIKVISKYPDKXDYTTIOQTVMQTTIN 568
D 487 LVDVTDEKKNMKTYSNCSVNLDFGDISAYVWNTKFOYTNSSEPLVYQMATLSSIGN 546
QY 569 EYTGFEFRASVDNTAFSTSGGOGGDLPEKTYIGDYWEDVDKDIQ----- 618
D 547 -----KSVSTGNMAGFTNNOSGAG--OEVYKIGNYWEDTKNKGVOELGEKGVNV 596
QY 619 -----NTN----- 621
D 537 TVTVFDNNTNKKVGEAVYKEDGSYLIPNLPGNDYRVESNLPKGYEVTPSKQNNELDS 656
QY 622 -----DN-----EXPLSNVLTTL-- 634
D 657 NGLSSVITVNGKDNLSADLGITPKRYNLGDIWEDTKNKGIDODEKISGVTYTLAKDN 716
QY 635 -----TYPDG----- 639
D 717 GNVLTVTTDADGKXKFTDLNNGNKYVEFTPEGYTPTVTSGDIEKDSNGLTTGVIN 776
QY 640 -----TSSV-----RT 646
D 777 GADNMTLDSGFYKTPKYNLGNVWEDTKNKGKODSTEGISGVTYTLKNGEVLQTTKT 836
QY 647 DEDGKYQDGLKNGLTYKITEFTPEGYTPTLKHSGTNPALDSEGS----- 692
D 837 DKDKGYQGTGLENG--TYVVEFTPEGYTPTVOGSGTIDSGIDSNGISTGVIKDKNDITD 895
QY 633 -----VW-----VTI-----NGQ----- 700
D 896 SGFYKPYNLGDIWEDTKNKGKODDEKISGVTYTLKDEHDKLKTVTTDENKGYOFT 955
QY 701 -----DDMTIDSGFYOTPKXS 716
D 956 DLNNGTYVEFTPEGYTPTVTSGNDTEKDSNGLTTGVIKDANMTLDSGFYKTPKXS 1015
QY 717 LGNYVWYDNKDGIOGDEDEKISGVYTLKDENGNIISTTTDENGKYQFDNLNSGNYIV 776
D 1016 LGDYVWYNSNKGDKODSTEGIKDKVYLLNKEGEVIGTITKTDENKXIRFDLDSGKRYV 1075
QY 777 HFDKSGMGTQTTSDGDDDEODAGEVYVITDHDIFSIDNGYVDE--SDSDSDSDSDS 835
D 1076 IFEXGTGLGTGNTTEDD--KDADGCEVYVITDHDFTLNGYEEFTSDSDSDSDSDS 1134
QY 836 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 895
D 1135 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1194

```

```

QY 896 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 955
D 1195 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1254
QY 956 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1015
D 1255 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1309
QY 1016 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1064
D 1310 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1359
QY 1065 KGTLLGLTFLAGLALL-GKRRKNRK 1089
D 1360 NATLFGCLFALGSLFLGRRKKOKK 1385

RESULT 9
QY 099M48 PRELIMINARY; PRT; 953 AA.
AC 099M48
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein.
GN SDRG OR SAV0561 OR SA0519.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iano J.-O., Ito T.,
RA Karamori M., Matsunaru H., Maruyama A., Murekami H., Hoshiyama A.,
RA Mitutani U.I., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hiraizumi A., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hirayama K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus ";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003359; BAB56723.1;
DR EMBL; AP003131; BAB41750.1;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;

Query Match 37.1%; Score 2093.5; DB 16; Length 953;
Best Local Similarity 43.8%; Pred. No. 2.8e-63;
Matches 488; Conservative 140; Mismatches 298; Indels 189; Gaps 28;
QY 2 INKNNLLTKRRKPIANKSNKAYAIRKFTVGTASIVGATLLFGLGHNKAEENSVDYKD 61
D 1 MNKKTATNRKGMPLNRKLSIRKYSVGTASILVGTLLFGLSGHKAKE----- 52
QY 62 SNTDELSDQSDSDEKNDVYNNQSIINDNNOIITKEETNNYDGIKRSDEPTEST 121
D 53 -HTNGLNQSNETTAPSEN-----KTT 74
QY 122 TNVDENEATFLQKTPQDFTHLTEEEVKSSSVSSNSIDTAQOPSHTTINREVSQSD 181
D 75 EKVDSROL-----KDNVQTA-----TADQPKVT----- 97
QY 182 NVEDSHVDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEK---ISNQ 238

```

```

Db 98 -----MSDSATVETSSMOS-----PQNAFASOSTQTSNVTINDKSSITYSNE 142
OY 239 DELNLPIREYEN-KARPLSTTSAOPSIRKRVYNOLA-EGSNNVHLIKVYDOSI----- 292
Db 143 TDKSNL--TQAKNVSTPTPTTITIKORALNRMAVNTVAALQOGTINVDKHFNINIDAIK 200
OY 293 -----TEGYDDESEGVAKAHADENLIYDVFEDVDKVKSGDMVTVDIDKNTPVSDLT 343
Db 201 GHVNTKTGTEFMATSSDYLK-----LKANYTIDDSVKEGDFTEFFKGYGFRGVSVR 252
OY 344 DPTTPKIDNSGEIATGTYNKKQITTYTDTYDKENIKAHKLKLSYIDKSKVPNN 403
Db 253 LPSQTONLNAOGNIIAKIYDSKTTTYYTFTNYVDQYTN-SGSEQVAFAKRENATTD 312
OY 404 NKLDEYETALSSVKKITTEVEORPENRNTANLQSMFNITKNTHTDQITYN-PLRY 462
Db 313 KTAIVKEVTLGNDYTSKDYIVDI--GNKGQOLISSTYINNEEDSRMNTVYVNPCKT 369
OY 463 SAKETNV-NISGNGDEGSTIIDSTIIKVKYKVDNONLPSHRIYDYSEVEDYND-DYA 520
Db 370 YTKETFTVTLT-----GYKFNPDAKNFKEYETDQNFVDS-FTPTPSKLDVTDGQFDVI 423
OY 521 QLGNNNDVINFGN---IDSPYIIVISKYDPNKDVTYITQOTVTMOTTINETYGEFRT 576
Db 424 YSNDKATATVDLNGOSSSDKOYIIQOVAVPNSSTDNCKIDYTTLETQNGKSSMSN---- 479
OY 577 ASYDNTIATSSGOGQLPREPKYKIGDYWEDVDKGIQNTNDNEKPLSVLTLTY 636
Db 480 -YSVSNVNGSSTANGD-----OKRYNLGDYWEDVNDKKGKDA--NEKGIGKVYILKD 529
OY 637 PDGTS-KSVRTEDGKYOPDGLKNGLYKITEFPEGYPTLKHSGTNPALDSEGSVWY 695
Db 530 SNGKELDRTTTDENGKYYOTGLSNG-TYSVEFSTPAGYIPTTANAGTDAVDSGLTTTG 588
OY 696 TINGODDMTIDSGFYOTPKYSLGNYWYDTNKKGIOGDDEK3ISGYKVTILKDENGNIST 755
Db 589 VIKDADNMILDSGFYKTPKYSILDYVWYNSKNGKODSTEEKIGKVKYVLQNEKEGYIGT 648
OY 756 TTTDENGKYOFOPLNSGNYIVHFDPKSGMOTTTDSGDDEDDADAGEVAVTTTDDHDS 815
Db 649 TETDENGKRFNDLDSGKRYVTEKPAGLTQGTNTTEDD-KPADGEVAVTTTDDHDT 707
OY 816 IDNGYVDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db 708 LDNGYEEET-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 766
OY 876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db 767 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 826
OY 936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db 827 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 886
OY 996 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1055
Db 887 DSDSDSDAKHPTPKPMSTV-----KDHKTAKALPE 918
OY 1056 TGAENYGSKGLTGLTFAGLGALL-GKRRKNRK 1089
Db 919 TGSENNNSNNGTLFGLFALGLLFGRRKKONK 953

```

```

RESULT 10
ID 086487 PRELIMINARY; PRT; 947 AA.
AC 086487;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE SARC protein.
GN SARC.

```

```

OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_Taxid-1280;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-NEMAN.
RX MEDLINE-99098700; PubMed-9884231;
RA Josefsson E., McCreia K., Ni Eldhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "three new members of the serine-aspartate repeat protein multigene
family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL; AJ005645; CAA0650.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor. 1.
DR TIGRFAMs; TIGR01167; LPTG_anchor. 1.
DR TIGRFAMs; TIGR01168; YSTRK_signal. 1.
DR PROSITE; P500343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 947 AA; 10288 MW; 36EFD6E35121534 CRC64;

Query Match 35.8%; Score 2023.5; DB 2; Length 947;
Best Local Similarity 42.3%; Pred. No. 6.2e-61;
Matches 470; Conservative 158; Mismatches 297; Indels 185; Gaps 28;

OY 2 INKNNLLTKRKPIANKSNKYAIRKFTVGTASIVGATLFLGLHNEAKAENSVDVXD 61
Db 1 MNKKATATNRKGMIPRLNKFISIRKYSVGTASIVGTTLIFGLSGHEAKAAE----- 52
OY 62 SNTDDELSNDSSDSEKENDVNNNQSINTDNNQITKEETNNVDGLEKREDTST 121
Db 53 -HTNGLNOSKNTTAPSEKTKKVDYSLQKNDTQ-----T 88
OY 122 TVNDENAEPLQKTPQDNTHTLFEVKESSSVSSSIDTAQPSHTTINRESVQTS 181
Db 89 ATADQPKVT-----MSDSATVETSSNM---QSPQNAFASOSTTQTSNV 129
OY 182 NVEDSHVDFANSKIKESNTSEKKEENTLEQPKVKVEDSTSQPSGTINDEKISMODEL 241
Db 130 TTNKDSSTYSNEDTSNLTQA-----KDVSTT----- 157
OY 242 LNLPIREYENKARPLSTTSAOPSIRKRVYNOLA-EGSNNVHLIKVY--DOSITEGY-D 297
Db 158 -----PKTTTTPKTLNRMAVNTVAALQOGTINVDKHFNINIDAIKGHVN 204
OY 298 DSEGVAK--AHDAENLIYDVFEDVDKVKSGDMVTVDIDKNTPVSDLTDSFTIPKIDNS 355
Db 205 QTGKTEFMATSSDYLKLNANYTIDDSVKEGDFTEFFKGYGFRGVSRLPSQTONLYNAQ 264
OY 356 GEIITAGTYNNKKQITTYTDTYDKENIKAHKLKLSYIDKSKVPNNNTKLDVEYKTL 415
Db 265 GNIIAKGIYSTNTTYYTFTNYVDQYTNVGSFEQVAFKRNKATTDKAYKKEVTLGN 324
OY 416 SSVNKTTEVEORPENRNTANLQSMFNITKNTHTDQITYN-PLRYSAKETNV-NISG 473
Db 325 DTVSEETIYDI--GNKKAQPLISSTYINNEEDSRMNTVYVNPCKATYTKQIFVTULT- 380
OY 474 NGDEGSTIIDSTIIKVKYKVDNONL-----PSNRITYDYSEVEDYNDYALQGNNDV 528
Db 381 -----GYKFNPDAKNFKEYETDQNFVDSFTPTPSKLDVTDGQFDVI--YSNDKATATV 433
OY 529 NINEGNIDS--PIIKYI-----SKYPNKDVTYITQOTVTMOTTINETYGEFRTASYDN 581
Db 434 DLKKGQTSNKQYIIQOVAVPNSSTDNQIDY-----TLDDKTKYSM--SNSYSN 483
OY 582 TIAFTSSGOGQDLPREPKYKIGDYWEDVDKGIQNTNDNEKPLSVLTLTYPGTS 641
Db 484 VNGSSTANGD-----OKRYNLGDYWEDVNDKKGKDA--NEKGIGKVYILKDSNGKE 534
OY 642 -KSVRTEDGKYOPDGLKNGLYKITEFPEGYPTLKHSGTNPALDSEGSVWVYTINGQ 700
Db 535 LDRTTTDENGKYYOTGLSNG-TYSVEFSTPAGYIPTTANAGTDAVDSGLTTGVYIKDA 593

```


[illegible]

```
Db 116 NETPOPTAIKNGATAKMODOTVPOEGNSQVD-----NKTNDANSTIATNSELKNSQTL 169
Qy 190 DFANSKIKESNTESGKEENTIEOPNKKVKEDSTTSQPSGYTNIDEKISNODELLNPINEY 249
Db 170 DLPOG-----SPQISNAQCTSKPSVTRAVRSLAVALPEVYN----- 206
Qy 250 ENKARPLSTSNQPSIKRVTYNQLAEOGSNNHLIKTDSITIEGICYDDSEGVIAHDAE 309
Db 207 -----AADAKGTNNV-----DKVTASNFLEKTEKTFPPNOSG 237
Qy 310 NLIYDFEVDOKVSGDMYV-----DID-----KNTVPGLDSTFIPIKONS 355
Db 238 NTFMANFTYVDKVSQDFTAKPLDPSLTGNGVDVYSSNNMTP-----IADKSTN 289
Qy 356 GEIATGTYDNKNKOITTYFTDYVDKYENIKAKHLKLTYSIDSKYPNNNTKLDVEKTLA 415
Db 290 GDVAKAKATYDILTKTYTFTDYVNNKKNKENGQFLPLFTDAKAPKSGT-YDANININAD 348
Qy 416 SSVNKTIVEYORP-----NENFTANLQSMFTNIDTKN-HIVEQTIYINPLRYSAKETNV 469
Db 349 EMFNKITYNASSPIAGIDKPNGANISSQILIGVDTASGQNTYKQTVFVNPVKORVLGNTVY 408
Qy 470 NISGNGDEGSTIIDDS-----TIIKYYKYGDNQNLPS-----N 503
Db 409 YIKGYQDK-----IESSGKVSATDTKLRIEYNDTSKLSDSYADPNDSNLKEVTDQENK 464
Qy 504 RIYDSEYEDVTNDDYAOGLGNNDYNINEGNIDSPYIIVIKISKYDPNKKDYTT--IQQTV 561
Db 465 RIY-----YE-----HPNVASIKRGDITKTYVVLVECHYDWTGKLNLTQYIOENV 509
Qy 552 TMOITINEYTGFEPTASIDNTIAFTSTSSGOGGDLPEKTYTIGDYWEDVDKQIGQNTN 621
Db 510 -----DEVTDNDYCI--FGW-----NN 524
Qy 622 DNEKLSNVLTLYTPDGSKSVRTDEDEGKYOPDGLKNGLTIKTFEPPEGYTPTLKHS 681
Db 535 EV-----VVRYGGSQA-----DG-----DSAVN-----PKDPTP----- 548
Qy 682 TNPALDSEGNISVWYTINGODDWTIDSGFYQPKYSIGANYWYDTNKGQIGQDEKISGV 741
Db 549 -GPPYDPEPSP-----DPEPEPTPDPPEPSPDPPEPSPDSDSDSDSGSD----- 595
Qy 742 KYTLKDENGNIISTTTTDENGKYQFDNLSNGYIVHFDKPSGNTOTTTDSGDDEODADG 801
Db 536 -----SDSGSDSDSESD- 607
Qy 802 BEVHYTIDHDPSIDNGYDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861
Db 608 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 660
Qy 862 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 921
Db 661 DSDSDSDSDSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 720
Qy 922 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 981
Db 721 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 780
Qy 982 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1041
Db 781 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 840
Qy 1042 G-----NSDCKSTR-----DKLPDTGANEDYGSKGTLLGLTFLAGLAL 1079
Db 841 RYTPPNNEQKAPSNPKGEVHNHNSKVKQHKTDALPETG-DKSENTNATLFGAMMALLSGL 899
Qy 1080 LL-GKRRKRRKK 1091
Db 900 LLEFRKKQDHKEK 912
```

result 15
Q932C5

```
ID 0932C5 PRELIMINARY: PRT: 935 AA.
AC Q932C5.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibriogen-binding protein.
CN FNB OR S4V0811.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Itoh T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshino A.,
RA Mizutani U.I., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hatakawa H., Kuhara S., Goto S., Iabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RL Lancet 357:1225-1240(2001).
DR EMBL, AP003360; BAB56973.1; -.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMS: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSRK_signal; 1.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;

Query Match 28.0%; Score 1581; DB 16; Length 935;
Best Local Similarity 36.7%; Pred. No. 4,4e-46;
Matches 412; Conservative 158; Mismatches 310; Indels 244; Gaps 36;

Qy 18 KSNKAIKRFYGTASIVGATLLFG-LGHNEAKAEENSVDQDKSMTDDELSDSDSS 76
Db 5 KKEKAIRKKSIGVASVLGTLIGGLSSKEADSENSV-----TQSDSASN 52
Qy 77 DEKNDVINNNOSINTDNNQIITKEETNNYQIEKRSDEKRESTTNDENATFLQKTP 136
Db 53 ESKSNDSSVSAAAPKTDPTDNV-----SDPKTSSTNN--NGETSAQAKP 93
Qy 137 QNTHTEEEVSSSVSSSSSIDTAQOPSHHTINREESVQTSQDVEDSHVSDFANSKI 196
Db 94 -----AQOETTQSSSTNATTEETPYTGATTTTNNQANTPATTS-SNTNABELVN--- 143
Qy 197 KESNTESGKEENTIEOPNKKVKEDSTTSQPSGYTNIDEKISNODELLN-PINEYENKARP 255
Db 144 QTSNETTSNDTNTV-----SSVNSPONSNTNABENASTQDTEATPSN---NESAP 191
Qy 256 LSTTSA-----QPSIKRVTYNQLA-----EGGSNNVHLIKVYTOOSTIEGIDSE 300
Db 192 QNTDASNKDVYQAVNPSPPRKRAFSLAVALADAPAAGTDITN--GLTVPKYT--IDSG 246
Qy 301 GYIKAHDAENLIYDYVFEVDOKVSGDTMTVIDDKTVPDSDLSFTIKIDNSGEIAT 360
Db 247 TTYVPHQAGYVKNLNFSPNSAVKCDTFTKITYPKLNLNGVSTYKAVPPINAGD-QVILA 305
Qy 361 TGTYDNKNKOITTYFTDYVDKYENIKAKHLKLTYSIDSKYVPNNNTKL-DVEKTLA--SS 417
Db 306 NGVIDS-DGNVLYITFTDYDNKKNENTANITMPAYID-----PENVTGTAVNTLTGIGTGN 360
Qy 418 VNKTIIVEYORPNNNTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNV-----N 470
Db 361 ASKTYLIDYEKGQFNLISIKGTIDQIDKTNNTYROTIIYVNP-----SGDNNVLPALTN 415
Qy 471 ISGNGDEGSTIIDSTIIVYKYGDNQNLPSNRIDYSEYEDVTNDDYAOGLGNNDVNI 530
Db 416 LIPNTKSNALIDAKNTDIAKRYV-DNANDLSESYIYNPSDFEDVTNQVATSPNANQYKV 474
```

```

OY 531 NF-----GNIDSPYIILKISKYDPNKNDDYTTIQOVTVMQOTINEYGEF--RFSYDNTIA 584
Db 475 EEPTEDDDDITTPYIYVNGHIDPASTG-----DLALSTFYGYDSNFIWRSMWNEVA 528
OY 585 FSTSGGOGGDLPEPKTKYIGDYWEDYDKGDIQONTNNEKPLSNVLVTLYPDGTSKSV 644
Db 529 FNNGGSGSDG-----IDKPYVP-----EQP----- 548
OY 645 RTDEPKYQFQGLKNGLYKIFETPEGYPTLKHSGTNPALDSEGNYSWVTINGQDDMT 704
Db 549 --bqGEIE-----PIPEDSDSDPGSDSGSDS-----NSDSG 578
OY 705 IDSGFYQTPKXSLGNVYWDYTNKDGIOGDDEKIGVAVTLKDENGNIISTTTDENGKY 764
Db 579 SDSGSDSTS-----DSGSDASDSD-----SASDSDSASDSDSASDSD 607
OY 765 QFDNLNSGNYIYHFKPFGMTQTTDSGDDDEQADGEEVHYTTDHDDFSIDNGYDDE 824
Db 608 DSDSASDSD-----SASDSDSASDSDSASDSD-----SASDSDSASDSDSASDSD 652
OY 825 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 884
Db 653 SASDSDSASDSDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 712
OY 885 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 944
Db 713 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 772
OY 945 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1004
Db 773 SDSDSESDSDSDSDSDSDSDSDSDSDSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 832
OY 1005 SDSDSDSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1047
Db 833 SGSDSDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 892
OY 1048 STDKLPTPGANEDYGSKGTLLGLFAGIGALLGKRRKRNK 1091
Db 893 DSKEPLPTGSEDE--ANTSLIWGLASIGSLLEFRKKENKDK 934

```

Search completed: March 14, 2003, 13:00:24
 Job time : 60.7491 secs

Result No	Score	Query Match	Length	DB	ID	Description
1	3013	100.0	1092	19	AAW41602	Staphylococcus epi
2	2810.5	93.3	930	23	ABE40469	Staphylococcus epi
3	2810.5	93.3	991	21	AAV83171	Cell wall protein
4	2810.5	93.3	991	21	AAV70120	Staph. epidermidis
5	1144.5	38.0	1166	20	AAV08643	S. aureus SdrE pro
6	857	28.4	278	22	AAO82803	S. epidermidis ope
7	576	19.1	1315	20	AAV08642	S. aureus SdrP pro
8	576	19.1	1349	22	AAV34402	Staphylococcus au
9	576	19.1	1349	22	AAU37544	Staphylococcus au
10	499.5	16.6	932	22	AAU34082	Staphylococcus au

	RESULT 1
AAW41602	ID AAW41602 standard; protein; 1092 AA.
XX	
AC	AAW41602;
DT	22-JUN-1998 (first entry)
DE	Staphylococcus epidermidis fibrinogen binding protein FIG
XX	
KW	Fibrinogen binding protein; FIG; aggregation; infection;
KM	coagulase-negative Staphylococcus; therapy; diagnosis;
KX	immunisation; immunogen; vaccine.
OS	Staphylococcus epidermidis strain HB.
XX	
FT	Key
FT	Peptide
FT	/label= Sig-peptide
FT	Protein
FT	/label= Mat-protein
FT	Region
FT	52..824
FT	/note= "non-repetitive region, harbours
FT	fibrinogen binding activity"
FT	Region
FT	825..1040
FT	/note= "asp-ser dipeptide repeat region";
FT	1053..1057
FT	/note= "cell wall anchoring motif"
XX	
PN	MO9748727-AI.
PD	24-DEC-1997.
XX	

PF 18-JUN-1997; 97WO-SE01091.
 XX 20-JUN-1996; 96SE-0002496.
 XX (FLOC/) FLOCK J.
 PA (FRYK/) FRYKBERG L.
 PA (GUSS/) GUSS B.
 PA (LIND/) LINDBERG M.
 PA (NILS/) NILSSON M.
 PI Flock J, Frykberg L, Guss B, Lindberg M, Nilsson M;
 XX MPI: 1998-063079/06.
 DR N-PSDB: AAV04279.
 XX
 PT Fibrinogen-binding protein from coagulase-negative Staphylococcus
 used for prevention, treatment and diagnosis of staphylococcus
 infection
 XX
 XX Example 3; Fig 6; 45pp; English.
 CC The protein comprises the fibrinogen binding protein (FIG) of
 CC coagulase-negative Staphylococcus epidermidis HB. Its amino
 CC acid sequence was deduced from the isolated fig gene (see AAV04279).
 CC The closest known analogue of FIG is the clumping factor of
 CC S. aureus which also binds fibrinogen and promotes bacterial
 CC aggregation in serum. Recombinant FIG polypeptides can be
 CC expressed in host cells. They are used as immunogens, particularly
 CC in vaccines (which may be expressed in vivo) to protect humans and
 CC animals against coagulase-negative Staphylococcus infection.
 CC Antibodies raised against FIG can be used for passive immunisation.
 CC They block the adherence of bacteria and for diagnosis.
 XX
 SQ Sequence 1092 AA;

Query Match 100.0%; Score 3013; DB 19; Length 1092;
 Best Local Similarity 100.0%; Pred. No. 1.2e-156;

Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDEKNDYINNNSQINTDDNNQIIKKEETNNYDIEKRSDETESTTNDENEATFLOK 60
 DB 75 SSDEKNDYINNNSQINTDDNNQIIKKEETNNYDIEKRSDETESTTNDENEATFLOK 134
 QY 61 TPQDNHLEEEYKSSSVSSNSIDTAQPSHTTINRESVQTSNDNEDSVSPFANS 120
 DB 135 TPQDNHLEEEYKSSSVSSNSIDTAQPSHTTINRESVQTSNDNEDSVSPFANS 194
 QY 121 KIKESNTESGKEENTIEOPKRVKEDSTISQPSGYTNIDEKISMODLLNPIREYENKAR 180
 DB 195 KIKESNTESGKEENTIEOPKRVKEDSTISQPSGYTNIDEKISMODLLNPIREYENKAR 254
 QY 181 PASTSAQSIRKVTYNOLAEOGSNVNHLIKYTDOSITEGYDSEGVAKHAENLIYD 240
 DB 255 PASTSAQSIRKVTYNOLAEOGSNVNHLIKYTDOSITEGYDSEGVAKHAENLIYD 314
 QY 241 VFEEVDKYSQGTMTVDIDKNTVPSDLTDSFTPIKIDNSGFIITATGYDNKKOITYT 300
 DB 315 VFEEVDKYSQGTMTVDIDKNTVPSDLTDSFTPIKIDNSGFIITATGYDNKKOITYT 374
 QY 301 FTDYVYKYEIKAKHLKTSYIDSKVPNNNTKLDVEKTAALSSVNTITTEYQRPENRT 360
 DB 375 FTDYVYKYEIKAKHLKTSYIDSKVPNNNTKLDVEKTAALSSVNTITTEYQRPENRT 434
 QY 361 ANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGDGSIIIDSTIIKKYKYG 420
 DB 435 ANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGDGSIIIDSTIIKKYKYG 494
 QY 421 DNQNLDSNRITVYSEEDYTDNDYQAQGNNDVYNINFGNIDSPYIIKYSKTDPKKDY 480
 DB 495 DNQNLDSNRITVYSEEDYTDNDYQAQGNNDVYNINFGNIDSPYIIKYSKTDPKKDY 554
 QY 481 TTIGQTVTQMTTINETGCFRTASYDNTIAFSTSSGCGGDDLPPEKTYIKGYVMEVDYK 540

DB 555 TTIGQTVTQMTTINETGCFRTASYDNTIAFSTSSGCGGDDLPPEKTYIKGYVMEVDYK 614
 QY 541 DGIQNTDNEKPLSNVLTLYTPDGTSKSVRTDEDEKRYGDFG 582
 DB 615 DGIQNTDNEKPLSNVLTLYTPDGTSKSVRTDEDEKRYGDFG 656

RESULT 2
 ABP40469
 ID ABP40469 standard; Protein; 930 AA.
 XX

AC ABP40469;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.

KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX antibacterial; gene therapy.

OS Staphylococcus epidermidis.

PN US6380370-B1.

PD 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

DR MPI: 2002-381255/41.

DR N-PSDB: ABN93014.

PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis

PS polypeptide, useful for diagnosing and treating bacterial infections -

XX Disclosure; SEQ ID 5314; 267pp; English.

CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading

CC frame (ORF) nucleic acid sequences which encode the amino acid sequences

CC given in ABP35124 to ABP37960. The S. epidermidis sequences have

CC antibacterial activity and can be used in gene therapy. The sequences

CC can also be used in the diagnosis and treatment of bacterial infections,

CC particularly S. epidermidis infections. The sequences can be used to

CC screen for compounds able to interfere with the S. epidermidis life

CC cycle or inhibit S. epidermidis infection.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

XX USPIO web site.

XX Sequence 930 AA;

QY Query Match 93.3%; Score 2810.5; DB 23; Length 930;
 Best Local Similarity 93.6%; Pred. No. 1.2e-145;
 Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEKNDYINNNSQINTDDNNQIIKKEETNNYDIEKRSDETESTTNDENEATFLOK 60
 DB 73 SSDEKNDYINNNSQINTDDNNQIIKKEETNNYDIEKRSDETESTTNDENEATFLOK 131
 QY 61 TPQDNHLEEEYKSSSVSSNSIDTAQPSHTTINRESVQTSNDNEDSVSPFANS 120
 DB 132 TPQDNHLEEEYKSSSVSSNSIDTAQPSHTTINRESVQTSNDNEDSVSPFANS 191
 QY 121 KIKESNTESGKEENTIEOPKRVKEDSTISQPSGYTNIDEKISMODLLNPIREYENKAR 180
 DB 192 KIKESNTESGKEENTIEOPKRVKEDSTISQPSGYTNIDEKISMODLLNPIREYENKAR 251

XX Staph. epidermidis serine-aspartate repeat region protein Sdrg.
 DE
 XX
 KW Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAM;
 KW microbial surface components recognising adhesive matrix molecules;
 KW collagen binding protein; CBP; CNA; fibrinogen binding protein;
 KW Clumping factor A; ClfA; Clumping factor B; ClfB; FBP;
 KW fibrinectin binding protein; Staphylococcus infection;
 KW serine-aspartate repeat region protein; SDR protein; Sdrg.
 XX
 OS Staphylococcus epidermidis.
 XX
 XX Key Location/Qualifiers
 FT Msc-difference 14 /note- "Encoded by in-frame stop codon TAG"
 FT Msc-difference 33 /note- "Encoded by in-frame stop codon TGA"
 FT Msc-difference 964 /note- "Encoded by in-frame stop codon TAA"
 FT Msc-difference 980 /note- "Encoded by in-frame stop codon TAA"
 FT Msc-difference 989 /note- "Encoded by in-frame stop codon TAG"
 FT Msc-difference 989 /note- "Encoded by in-frame stop codon TAA"
 XX WO200012131-A1.
 XX 09-MAR-2000.
 XX 31-AUG-1999; 99MO-US19727.
 XX 31-AUG-1998; 98US-0098439.
 XX
 PA (INH1-) INHIBITEX INC.
 PA (TEXA-) UNIV TEXAS A & M SYSTEM.
 PA (QUE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX
 PI Patti JM, Foster TJ, Hook M;
 XX
 DR WPI: 2000-237781/20.
 DR N-PSDB; AA251202.
 XX
 PT Composition used for generating immune response or for inhibiting
 PT microbial colonization in an animal comprises antibodies that bind
 PT collagen binding protein, fibrinogen binding protein and, optionally,
 PT fibronectin binding protein.
 XX
 YX Claim 8; Fig 4; 115pp; English.

CC The patent discloses multicomponent vaccines containing selected
 CC combinations of bacterial binding proteins termed MSCRAM (microbial
 CC surface components recognising adhesive matrix molecules) or their
 CC antipodides. A vaccine composition is provided that includes collagen
 CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein
 CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB),
 CC and optionally a fibronectin binding protein e.g. FBP-A.
 CC The vaccines are useful for imparting protection against a broad
 CC spectrum of Staphylococcal strains and for inhibiting microbial
 CC colonisation, especially of Staphylococcus aureus, in an animal.
 CC The combinations can also be used to select donor blood pools for the
 CC preparation of purified blood products for passive immunisation.
 CC The present sequence is a serine-aspartate repeat region
 CC protein, Sdrg from Staphylococcus epidermidis. The Sdr protein is
 CC useful in vaccine preparation in combination with specific
 CC bacterial binding proteins. These vaccines can be used to treat a broad
 CC spectrum of bacterial infections, including those arising from both
 CC coagulase-positive and coagulase-negative bacteria.
 CC
 XX Sequence 991 AA:

Query Match 93.38; Score 2810.5; DB 21; Length 991;
 Best Local Similarity 93.68; Pred. No. 1.3e-145;
 Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEKNDVNNOSINTDDNNQIIKKETNYDIEKRSRDESTTNDENAEFLQK 60
 DB 106 SSNEKNDVNNOSINTDDNO-IRKEETNSDAIENRSKDLTOSTNDENAEFLQK 164
 QY 61 TPQDNTLITEEVSESSVSSNSIDTAAQPSHTTINRESVOTSHVEDSHSDPANS 120
 DB 165 TPQDNTLKEEVVAPESVSSNSMPTAAQPSHTTINSEASITSDNEERSRSDPANS 224
 QY 121 KIKESNTEGKEENTIQPNKVKEDSTTSGPGYTNIDEKISNDELINLPINEYENKAR 180
 DB 225 KIIESNTEGKEENTIQPNKVKEDSTTSGPGYTNIDEKISNDELINLPINEYENKAR 284
 QY 181 PLSTTSAOPSKRYTVNQLAAEGSNVNHILIKVQDSITTEGYDSDGIRAHDAENLIYD 240
 DB 285 PLSTTSAOPSKRYTVNQLAAEGSNVNHILIKVQDSITTEGYDSDGIRAHDAENLIYD 344
 QY 241 VTEFVDKVKSGDTMTYDIDKNTVPSDLTDSFTIPKIKDNGSELIANGTYDNKKQTYT 300
 DB 345 VTEFVDKVKSGDTMTYDIDKNTVPSDLTDSFALPKIKDNGSELIANGTYDNKKQTYT 404
 QY 301 FTDYVDKYEENIKAKLTLSTYIDKSKVPNNNTKLDVEYKTLASSVYKTIIVEYQKPNERT 360
 DB 405 FTDYVDKYEENIKAKLTLSTYIDKSKVPNNNTKLDVEYKTLASSVYKTIIVEYQKPNERT 464
 QY 361 ANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDESGTIIIDSTIIKVKYVG 420
 DB 465 ANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDESGTIIIDSTIIKVKYVG 524
 QY 421 DNQNLPSNRITDYSEVEDYNDYDQALGNNDVNIENFGNIDSPYIIKIVISKYDPNKKDY 480
 DB 525 DNQNLPSNRITDYSEVEDYNDYDQALGNNDVNIENFGNIDSPYIIKIVISKYDPNKKDY 584
 QY 481 TTIOQTNTMOTTINIEYGEFRTASYDNTIAFSTSSGOGGDLPEPKYKIGDYWEDEVK 540
 DB 585 TTIOQTNTMOTTINIEYGEFRTASYDNTIAFSTSSGOGGDLPEPKYKIGDYWEDEVK 644
 QY 541 DGIONTNDNEKPLSNVLTLYTPGTSKSVRTDDEGKYPFDG 582
 DB 645 DGIONTNDNEKPLSNVLTLYTPGTSKSVRTDDEGKYPFDG 686

RESULT 5
 AAY08643
 ID AAY08643 standard; protein; 1166 AA.
 XX
 AC AAY08643;
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE S. aureus Sdrg protein.
 XX
 KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; Sdrg; Sdrg;
 KW Sdrg; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
 KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
 KW extracellular matrix; vascular graft; vascular stent; vaccine;
 KW intravenous catheter; artificial heart valve; cardiac assist device;
 KW antibacterial.
 XX
 OS Staphylococcus aureus.
 XX
 PN MO9927109-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 25-NOV-1998; 98MO-US25246.
 XX
 PR 31-AUG-1998; 98US-0098427.
 XX
 PR 26-NOV-1997; 97US-0066815.
 XX
 XX (EIDH/) EIDHIN D N.
 PA (FORF-) FORAS T/A BIORESEARCH IRELAND.
 PA (FOST/) FOSTER T J.
 PA (HOOK/) HOOK M A O.

PA (INHIT-) INHIBITEX INC.
PA (JOSE/) JOSEFSSON E.
PA (PAT/) PATTI J M.
PA (PERK/) PERKINS S E.
XX
XX
PI Eldham DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
PI Perkins SE;
XX
XX
DR WPI; 1999-357844/30.
DR N-PSDB; AAX77594.
XX
XX
PT Staphylococcus aureus fibrinogen-binding proteins for treating
PT septicemia, osteomyelitis, mastitis or endocarditis
XX
XX Claim 8; Fig 9; 143pp; English.

Sequence 1166 AA;

Query Match	38.0%	Score 114.5	D3 20	Length 1166
Best Local Similarity	42.6%	Pred. No. 2e-54		
Matches 263	Conservative 108	Mismatches 191	Indels 55	Gaps 18

```

QY      1  SDEEKNDVYNNNOSINTDDNNQI--LKKEETNNYDQIEKRSEDRIESTNNVDENAEFL 58
Dh      66  ATTSNKEVEVETENNSTENNSTNNPIKKE--TNTDQBPAAKEESTSSSTOKOONVYTAJ 123
QY      59  OKTPODNHLNBEEEKES---SVSESSNSSIDTNOQPSHTTINNEEVOTSDNVEDSHVS 115
Db      124  TETKQON--IKENVKPSTDKTAEDTSVILEEKKAPNT--NNDVYTKS-----TS 172
QY      116  DFANSIKESNT---ESGKRENTIEOPNKYKEDSTSQSPSGNTIDEXIS--NODELNLP 171
Dh      173  EPSTSEIGTKFTPTTQESINENISQOPQPPKVD--NQVYDATNPKRPVNVSEELKNP 229
QY      172  -----INDEYENKARPLSTTSNOQSIRKY-----TYNQLAEOGSSVNVHLKYTDQ 216
Dh      230  EKLKELVBNDSNTDSTPVPATPVSAPKRNKMRFAVAOAPAAVSNVNNLDIKYTKQ 289
QY      217  STIEEYDSEGVYKAHDKENLIYVTFEVDVKKSGGPMVWDIDKNVPSLDTSFPIK 276
Dh      290  TIKVG-DGKDNVAAAHDKDDEIETDEFTIDKNVAKGDMTINDKNVIPSOLDKNDPID 348
QY      277  IKDNGEIIANGTYONKRAKQOITVFTDVDVDEYENIKAHKLTYSIDSKVPPNNNTKLDE 356
Dh      349  IDPSGEVIYAGTFKAPKQOITYTFTVDYDKEDIKSRLFLYSYIDKVP--NETSLNT 407
QY      337  YKTALLSSVNTITVEYQAPNENRANALOSMTNIDTKNHEVQOITLINPLRSKKEYNVN 396
Dh      408  FTAACKESONVATYODPBMVHGSSNDSISITTKLDEKQOTEOOIVYNPLKKSSTNTKVD 467

```

[illegible]

RESULT 6

ID AAG82803 standard; Protein; 278 AA.

AC AAG82803;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2700.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis;

KW vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN W0200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimmery WJ;

DR WPI; 2001-316495/33.

XX

PT useful for vaccinating against infections, e.g. endocarditis -

PS Claim 18; Page 705; 2188pp; Engl11sh

AAH52304 to AAH52970 represent nucleic acids (I) encoding polypeptides (II), given in AAH81454 to AAH83120, from *Staphylococcus epidermidis* (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53721 to AAH5090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 278 AA;
Query Match 28.4%; Score 857; DB 22; Length 278;
Best Local Similarity 93.0%; Pred. No. 1.6e-39;
Matches 172; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 SDEKNDVNNNSINTDNNQIKKEETNNYDGIKRSDEKREESTNVDENEATFLOK 60
DB 75 SDEENOVNNNSINDNNQIKKEETNNNSDIEKSEKREESTNVDENEATFLOK 134
QY 61 TPQDNTHTLEEVKSSSVSSSSSIDTAQPSHTTINRESVQTSNVEDSHVSDFANS 120
DB 135 SPQDNTHTLEEVKSSSVSSSSSIDTAQPSHTTINRESVQTSNVEDSHVSDFANS 194
QY 121 KIKESNTSGKEENTIEOPNKVKEDSTTSOPSGYNTIDEKISNODELNLPIENENKAR 180
DB 195 KIKESNTSGKEENTIEOPNKVKEDSTTSOPSGYNTIDEKISNODELNLPIENENKAR 252
DB 181 PLSTP 185
DB 253 KLSVT 257

RESULT 7
AY08642
ID AY08642 standard; Protein: 1315 AA.
XX
AC AY08642;
XX
DT 09-AUG-1999 (first entry)
XX
DE S. aureus SdrD protein.
XX
KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;
KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
KW extracellular matrix; vascular graft; vascular stent; vaccine;
KW intravenous catheter; artificial heart valve; cardiac assist device;
KW antibacterial.
XX
OS Staphylococcus aureus.
XX
PN MO9927109-A2.
XX
PD 03-JUN-1999.
XX
VX 25-NOV-1998; 98MO-US25246.
PR 31-AUG-1998; 98US-0098427.
PR 26-NOV-1997; 97US-0066815.
XX
PA (EIDH/) EIDHIN D N.
PA (FORF-) FORFAS T/A BIORESEARCH IRELAND.
PA (FOST/) FOSTER T J.
PA (HOOK/) HOOK M A O.
PA (INH-) INHIBITEX INC.
PA (JOSE/) JOSESSON E.
PA (PAT/) PATTI J M.
PA (PERK/) PERKINS S E.
XX
PI Elphin DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
PI Perkins SE;
XX
DR WPI: 1999-357844/30.
DR N-PSDB: AAX77593.
XX
PT Staphylococcus aureus fibrinogen-binding proteins for treating
XX septicemia, osteomyelitis, mastitis or endocarditis
XX
PS Claim 8; Fig 8; 143pp; English.
CC This invention describes novel Staphylococcus aureus fibrinogen-binding
CC proteins that bind both the alpha and beta fibrinogen chains. The

CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and
CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere
CC to medical devices, binding proteins that bind both the alpha and beta
CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as
CC competitive inhibitors to block this binding. Antibodies against ClfB,
CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
CC The proteins of the invention can be used in a pharmaceutical composition
CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
CC S. aureus to the extracellular matrix. The proteins or their fragments
CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than ClfA), S. hemolyticus, S. lugdenensis, and S. schleiferi.
CC The proteins of the invention have antibacterial activity.

Sequence 1315 AA;
Query Match 19.1%; Score 576; DB 20; Length 1315;
Best Local Similarity 30.1%; Pred. No. 2.9e-23;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 1 SDEKNDV---INNNSINTDNNQIKKEETNNYDGIKRSDEKREESTNVDE-NEAT 56
DB 55 STNKLNEATTSASDNDQSDVDMQQLNEDNTKNDQKEVSSQGNETTNGKMLIKE 114
QY 57 FLOKTPQDNTHTLEEVKSSSVSSSSSIDTAQPSHTTINRESVQTSNVEDSHSD 116
DB 115 SVQSTGKVKVEVSTAKSDEQASPKSTNEDLNTKQ---TISNDEALQ-PDLQEKSVVN 168
QY 117 FANSKIKESNTSGKEENTIEOPNKVKEDSTTSOPSGYNTI-DEKISNODEL--NLPIN 173
DB 169 ---VOPTNEENKKVD-----AKTESIT-----LWKSDAIKENDELVDLNNNSNS 209
QY 174 EYENKARPLSTSAOP---SIKRVYVNOQLAEGGSNVNHLIKYVDOSITEGYDSECVI 229
DB 210 NNEENNADILILPKSTAPKRLNTRMRAIAVQSPSTKAKNNDLITSNTYLVVADAKNKKIV 269
QY 230 KAHDAENILYVTFEVDKVKSGDTMTVDIDKNY-----PSDLTDSFTLPKIKD-NSG 282
DB 270 PAQDYLSLKSQIT--VDDKVKSGDYFTIKY--SDTYQVYGLNPEIDKN--IGDIRKDPNNG 323
QY 283 ELIATGTDNKKNKOLTYFTDYVDKYENIKAKHLKLSYIDKSKYPNNNTKLDVEYKATLS 342
DB 324 ETITAKHDTANMLITTYFTDYVDRFNSVOMGINKSITMDATIP--VSKNDVEFNVTIG 381
QY 343 SVNKTITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTIYINPLRYSAKE 392
DB 382 NTTTKRTANIQYDPVYVNEKNSIGSAFT--ETVSHVGNKKNPGYKQIIVYVNSPNSLTN 439
QY 393 TNNVNI-----SNGDEGSTIIDDSITIKYKKGDNONLPSNRIDYS--EDEDVTINDYA 446
DB 440 AKLVQAVHSSYPNNIGQINKVDIKIYQVKGTYL--NKGYDVNKKELTDVYN-OYL 495
QY 447 Q---LGNNDVNIENGNDSPYIIRKIVISKYDPNKDQYTTIQYTYMOTINEYGEFRTA 503
DB 496 OKITYGDNNSAVIDFGNADSAVYVWNTKRFQYTSSESTLVOMATLSSTGN-----KSV 549
QY 504 SYDNTIASTSSGOGGDLPEPKTYKIDYVWEDVDKDGIGONTDNEKPLSNVLTLYTP 563
DB 550 STGNALGFTNNQSGAG---QEVYKIGINYMEDPNKNGVQEL--GEKGVGNVTVTV-FD 602
QY 564 DGTSKSVR---TDEDGKY 578
DB 603 NNTNTKVGGAATKEDGSY 620

RESULT 8

```

AAU34402
ID  ID  AAU34402 standard; Protein: 1349 AA.
XX
AC  AAU34402:
XX
DT  14-FEB-2002 (first entry)
XX
DE  Staphylococcus aureus cellular proliferation protein #678.
XX
KW  Antisense; prokaryotic cellular proliferation protein;
KM  antibiotic; antibacterial; drug design.
XX
XX  Staphylococcus aureus.
OS
WO200170955-A2.
XX
PN  27-SEP-2001.
XX
PD  21-MAR-2001; 2001WO-US09180.
XX
PF  21-MAR-2001; 2000US-191078P.
XX
PR  21-MAR-2001; 2000US-191078P.
XX
PR  23-MAY-2000; 2000US-206848P.
XX
R  26-MAY-2000; 2000US-207727P.
XX
R  23-OCT-2000; 2000US-242578P.
XX
PR  27-NOV-2000; 2000US-253625P.
XX
PR  22-DEC-2000; 2000US-257931P.
XX
PR  16-FEB-2001; 2001US-269308P.
XX
PA  (ELIT-) ELITRA PHARM INC.
XX
PI  Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT;
PI  Yamamoto RT, Xu HH;
XX
XX  MPI: 2001-611495/70.
DR  N-PSDB: AAS52261.
XX
XX  New polynucleotides for the identification and development of
PT  antibiotics, comprise sequences of antisense nucleic acids -
XX
XX  Example 3; Seq ID No 5898; 511pp; English.
XX
XX  The invention relates to antisense inhibitors of genes essential to
XX  prokaryotic cellular proliferation, their use in identifying the
XX  genes, their use in the discovery of novel antibiotics, the essential
XX  genes themselves and the encoded proteins. The prokaryotes used are
XX  Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX  pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX  invention is also useful for the identification of potential new targets
XX  for antibiotic development. The antisense nucleic acids can also be used
XX  to identify proteins used in proliferation, to express these proteins,
XX  and to obtain antibodies capable of binding to the expressed proteins.
XX  The proteins can be used to screen compounds in rational drug discovery
XX  programmes. The antisense nucleic acid sequence is also useful to screen
XX  for homologous nucleic acids which are required for cell proliferation in
XX  a wide variety of organisms. The present sequence represents an
XX  essential prokaryotic cellular proliferation protein.
XX  Note: The sequence data for this patent did not form part
XX  of the printed specification, but was obtained in electronic
XX  format directly from WIPO at
XX  ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 1349 AA;

Query Match 19.1%; Score 576; DB 22; Length 1349;
Best Local Similarity 30.1%; Pred. No. 3e-23;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 299

1  SSDEKNDV---INNQSINTDNNNOIITKEEFNNVYDGIKRSRDPDRSTNTNVE-NEAT 56
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
db  55  STNKEINERTTSASDQSSDDKDYDMQJNGBDNTKNDQKQKMYVSQGNHTTSNGKLLIEK 114
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57  FLKTPQDDTHTLIEEVKSSSVSSNSSIDTAQOPSHTTINREESVOTSDNVEDSHVSD 116
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Dh	115	SVQSTGTGKKVVEVSTAKSDEQASPRSTNEDLNTKQ-----TISNQEALQ-PDLQENKSVVN	168
Qy	117	FANSKIKESNTSECKEENTIEQPKVKVEDSTTSOPSGYTNI--DEKISNODELL--NLPIN	173
Dh	169	-----VOPTNENKKVVD-----AKTESTP-----LVNKSDAIKSNDETLVDNNNSN	209
Qy	174	EYENKARKPLSTTSAPQ-----SIKEVTVNQLAAEAGSVNVNHLIKYTKDQSLIEGVDDSGVI	229
Dh	210	NNENNADIIIPKSTAPRKLNTRMRIIAAVQSPSSTEAKKNVNDLTSNTLLTVVADADKNNKIY	266
Qy	230	KAHDAENLIVYTFEEVDVKGSGDPMTVDIDKNTV-----PSDLADSFTIPIKID-NSG	283
Dh	270	PAQDYLSLKSQIT--VDDKVASGGGYFLIKY-SDTVOYVGLNPEDIKN---IGDKIDPNNG	323
Qy	283	EIIATGTVDNKNKOITVTFDIDYDKYENIKAKHLKLTGYIDKSKVPNNNTKLDEVEYKTALS	342
Dh	324	ETIATAKHDFTANNLITVTFDIDYDRFNSVQMGINYSIYMDADTIP--VSKNDVEFNVTIG	381
Qy	343	SVNKTIYVEYRPN--EKRITANLQSMFTNIDTKNHT-----VQGTIYNPLRYSAKE	392
Dh	382	NTTKRTTANIDYPPYVNVKENSISAFY--ETVSHVGNKKNPGYKKQTIYVNPSENSLTJN	439
Qy	393	TNVNVI-----SGNGDEGSTIIDSTITIKYKKGDNQNLPSDRIRYDYS--EYEDVTNDYA	448
Dh	440	AKLVQAVHSSYIPNNIGQINKDVTDIKIQVPRKGYTL---NKGIDVATKELTDVTN-QYL	495
Qy	447	Q---LGNNDNVNINEGNIDSPYIIKIVISKYDPRNKDDYTTIOQTVMQTTINETYGEFRTA	503
Dh	496	QKITGQDNNSAVIDFGNADSAVYVMVNTKPKGYTNSSEPTLVQMATLSSTGN-----KSV	549
Qy	504	SYDNTIAFSTSSGGQGDLPPEKTYKIGDIYVMEVDKADGIONTDNENKPLSNVLTLYTYP	563
Dh	550	STGNALGTTNNQSGAG---QEVYKIGINYMEDTNKNGVEL--GEKGVGNVTVTV-FD	602
Qy	564	DGTSKSVR---TDEDGY	578
Dh	603	NNNTNKVGEAVTKEDGST	620
RESULT 9			
AAU37544			
ID	AAU37544	standard; Protein: 1349 AA.	
XX	AAU37544;		
AC			
XX			
DT	14-FEB-2002	(first entry)	
XX			
XX	Staphylococcus aureus cellular proliferation protein #1714.		
KW	Antisense; prokaryotic cellular proliferation protein;		
KM	antibiotic; antibacterial; drug design.		
XX			
OS	Staphylococcus aureus.		
XX			
PN	MO200170955-A2.		
PD	27-SEP-2001.		
XX			
PF	21-MAR-2001; 2001MO-USO9180.		
XX			
PR	21-MAR-2000; 2000US-191078P.		
PR	23-MAY-2000; 2000US-206848P.		
PR	26-MAY-2000; 2000US-207727P.		
PR	23-OCT-2000; 2000US-242578P.		
PR	27-NOV-2000; 2000US-253625P.		
PR	22-DEC-2000; 2000US-257931P.		
PR	16-FEB-2001; 2001US-269308P.		
XX			
PA	(ELIT-) ELITRA PHARM INC.		
XX			
PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ		
PI	Yamamoto RT, Xu HH;		
XX			

[illegible]

```

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
P1 Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR N-PSTDB: AAS54704.
XX
PT New polynucleotides for the identification and development of
PS antibiotics, comprise sequences of antisense nucleic acids -
XX Example 3; Seq ID No 12438; SIlpp: English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
SQ ftp.wipo.int/pub/published_pct_sequences.
Sequence 932 AA;
Query Match 16.6%; Score 499.5; DB 22; Length 932;
Best Local Similarity 27.7%, Pred. No. 2,8e+19; Indels 11; Gaps 24
Matches 160; Conservative 95; Mismatches 212;
QY 35 GIERSKSEDTTESTVNDENEATFLTKPDNHNLFEEVKSSSVSSSIDTAQPESH 94
Db | : : : : : ||| :||| :||| :||| :||| :||| :||| :||| :|||
45 GEAKAAEHTTNELNOSKHETT-----APSEN-----KTIEKKDSRQ---- 82
QY TTINREESYOVSITSDNEHVSPANSKIRESNTSGKEENTTEOPNKVKEDTTSOPLY 154
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
83 -----NNIQESTSD--DPKVNESDWTYSVE---TTEEP---QNTTSTOPTRKO 122
QY 155 TWIDEKISNODELNLPIETENKARPLSTSAPSIR-----RVYNOLLA-EQGSAVN 208
Db || :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
123 NN-DAMARKD--MLAONISTOAADVSTTPKRTTIKPRLNRMAVNVVAAPDOGTNYN 177
QY 209 HLIKVTGSI-----TEKYDSEGVIKAHDAENLIVDFEFVDOKVRKGDM 255
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
178 DKVFHTNIDIALDKGVNKTGTGEWARTSVDLK-----LRKNYTIDSVRKGDIF 229
QY 256 TVWDIKNPVPSDLTFSTIPRIKDNGSELLAGTYDNKKKQTITFTDYDKYEINKAH 315
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
230 TERGGGYFRPGSRVLPEQSQNLYNAOGNIIAKGIYDESETTFFFTFNVDQYTNIGSF 289
QY 316 KLTSLYIDRSKVPINNKLDEVKRYTALSANNKITTVERQRPNERNRANLOSMFNTDKRN 375
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
290 EOVAFKRRENATTDKAYLPFEVTLGNDRKSKNIVDY--GNOKGOOLLSTSNVINNEL 346
QY 376 IWEOTIYLIN-PLRYSAKETNV-NISGNDEGSTIIDSTIIKYYKRGDON-----LPD 427
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
347 SRKMITYAVNQPKTTYIKREFVTVLT-----GYFNPDDANKFKIYEV-TNQNGCVDSFTFD 400
QY SNRIYDYSEXEDVTNDYAQLGNNDVNINFG--NIDSPIILIKVISKYPDKQDYTTIOQ 485
Db :::: :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
401 TSKITLVDTVKRKFIT--SYNDNKTATVDLLINGOSSSDKQYIIIOVVAPPNMSSTDCKRIDY 457
QY TYVMOTTINEYGCEFFTASYDNITASTSSGCGQGLDLPRKPYKIJIDDYVEDVDKDG 545
Db :::: :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
486 TYVMOTTINEYGCEFFTASYDNITASTSSGCGQGLDLPRKPYKIJIDDYVEDVDKDG 545
Db :::: :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
458 TLETONGKSSWSN-----STSNVNSGSSTA GD-----OKRNLDGDIYWEDINDKROKOD 505
```

QY 546 TNDNEKPLSNVLYLTYPDGTG-KSVTDEDGKYQFDG 582
 DB 506 A--NEKIGKGVYILKDSNGEKELDRITTTDENGKYQFTG 541

RESULT 12

AY83170
 ID AAY83170 standard; Protein; 1802 AA.

AC AAY83170;

DT 24-JUL-2000 (first entry)

XX Cell wall protein SdrF.

XX SdrF: SdrF: coagulase negative; staphylococcus; septicemia;
 KW osteomyelitis; endocarditis; immune response; vaccine; graft;
 stent; intravenous catheter; heart valve; cardiac.

XX Staphylococcus sp.

XX Key Location/Qualifiers

FT MISC-difference 12 /note- "Position encoded by TAA stop codon"

FT MISC-difference 28 /note- "Position encoded by TAA stop codon"

FT MISC-difference 1771 /note- "Position encoded by TAG stop codon"

FT MISC-difference 1774 /note- "Position encoded by TGA stop codon"

FT MISC-difference 1793 /note- "Position encoded by TGA stop codon"

FT MISC-difference 1800 /note- "Position encoded by TAA stop codon"

XX MO200012689-A1.

PD 09-MAR-2000.

PF 31-AUG-1999; 99MO-0519728.

XX 31-AUG-1998; 98US-0098443.

PR 25-JAN-1999; 99US-0117119.

XX (QUEEN-) QUEEN ELIZABETH COLLEGE DUBLIN.
 (TEXA) UNIV TEXAS A & M SYSTEM.

XX Foster TJ, Hook M, Davis S, Hartford O, Mccrea K, Ni Eidin D;

DR WPI; 2000-256637/22.

DR N-PDB; AA293333.

XX Recombinant or synthetic proteins from coagulase-negative staphylococci
 PT useful for prevention, treatment and diagnosis of staphylococcal
 PT infections bind soluble and immobilized fibrinogen

XX Claim 2; Figure 2; 104pp; English.

XX Isolated staphylococcus sdr cell wall proteins which bind both
 CC soluble and immobilized fibrinogen are useful for treating or
 CC preventing coagulase-negative staphylococcal infection such as
 CC septicemia, osteomyelitis or endocarditis, and for inducing immune
 CC responses in patients. The cell wall proteins are also useful for
 CC reducing coagulase-negative staphylococci infection of indwelling
 CC medical devices such as vascular grafts, vascular stents,
 CC intravenous catheters, artificial heart valves and cardiac assist
 CC devices. The cell wall associated proteins are able to inhibit
 CC staphylococcal adhesion to immobilised extracellular matrix or host
 CC cells present on the surface of implanted biomaterials.

XX Sequence 1802 AA;

Query Match 16.6%; Score 499; DB 21; Length 1802;

Best Local Similarity 27.18; Pred. No. 7.2e-19;
 Matches 168; Conservative 104; Mismatches 275; Indels 72; Gaps 25;

QY 1 SDDEKNDVNNQISITDDNNQIIRKKEETNNYDGIERSEDESTETNDENEA-TFLQ 59
 DB 185 TSTTQODSTKKNPSLNDNLNSSSTTSKESKTDHSTKQAMST-NKSNLNDNPSPOSE 243
 QY 60 KT-PQDNTLHTEEVKSESSVESSNS-----SIDTAQOPSHTTIN-REESVQTSN 108
 DB 244 KTSQANNNDSDNQASPKDLSKPSQKVKYKTFKFNDEPTQDVHTTKLKLTPSVSDSS 303
 QY 109 VEDSHVDFANSKIKESNTSGKEENTIEQPNKKEEDSTSQPSGYINIDEKI----- 161
 DB 304 VNBK--QDYTRSAV---ASLGVDSNETETITNAVRONLDLKAASREQIEAIIAELAK 357
 QY 162 --SNODELLNP--INEYENKARPLSTTSAPSKRTVNOALAE--QGSVNHILRYTD 215
 DB 358 DFNMPDYGVTPLALNNSQKNSP--HKSASP--RNLMSLAEPNSGKRVNDRKAITN 412
 QY 216 Q--SITEGYDSEGVIRAKDAENLIYDTPREVDDKVRSGDTMTVDIDKNVPSDLTDSFT 273
 DB 413 PTLSLKSNHANNVIMPTSEQENLKANVELDSDIKEGDFTIKYQYIRPGLELPAI 472
 QY 274 IPKIKDNGSELIAFGTDNKNKQITTFDYVKNENIKAKLKTSLIDSKVNNNTKL 333
 DB 473 KTLRSKDGSIIVANGVYDKTNTTTFYFNVDQYQNTSGFDLIAIPKRETAIKDNQY 532
 QY 334 DVEKRTALSSVNTKIITYEORPENRPTANLQSMFTNIDTKHVTVEQTIYINPLASAKET 393
 DB 533 PMEYTTINEVYKADFIDYGNKKDNTT---AAYANDVNNKKNNEYVYLLNQNQNRKA 589
 QY 394 NVNISNGDEGSLIIDST--IKVYKVDNQLPDS-NRIYDSEVEDYTNDDYAO- 448
 DB 590 KY-----FSYVKNGEFIPEGVKVEYEDTNAMVDSFNDLNSVWKDVTSGFAPKVS 641
 QY 449 GNNNDVINFGNT--DSPYIKIYISKYDPKDKDYTIQCTVYQVTQITNETGFEFRASY 505
 DB 642 ADGTRVDINFRSMANGKKTIVQAVRPTGTGNYT--EYWLTRDGTIN--TNFYGRTK 697
 QY 506 DNTIAFSTSSGOGGDLPPERTYKIGDYVEDVDKDCIQNTDNEKPLSNVLYLTYPDG 565
 DB 698 STTYTYLNGSSTAGDNP---TYSLDGYVWLDRKKNQVQ--DDDEKGLAGYVYTL--KDS 750
 QY 566 TSKS---VRTDEDGKYQFD 581
 DB 751 NNRELQRTTDDSGHYQFD 769

RESULT 13
 AAY70119
 ID AAY70119 standard; Protein; 1802 AA.
 AC AAY70119;
 XX 06-JUN-2000 (first entry)

XX Staph. epidermidis serine-aspartate repeat region protein SdrF.
 DE Staph. epidermidis serine-aspartate repeat region protein SdrF.
 XX Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;
 KW microbial surface components recognising adhesive matrix molecules;
 KW collagen binding protein; CBP; CNA; fibrinogen binding protein;
 KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;
 KW fibronectin binding protein; Staphylococcus infection;
 KW serine-aspartate repeat region protein; SDR protein; SdrF.

XX Staphylococcus epidermidis.
 OS
 XX Key Location/Qualifiers
 FH MISC-difference 12 /note- "Encoded by in-frame stop codon TAA"
 FT MISC-difference 28 /note- "Encoded by in-frame stop codon TAA"
 FT MISC-difference 1771


```
OY 257 VDIKNTVPSDDLDSFTIPKIKDNGSELIATGTYDNKNKQITVTPFDVVKYENIKAHK 316
Db 240 FKYGQYFRPGSVRLPSOTONLYNAQNIIAKGIYDSTNTNTTYTFTNVDOYTNVRSFE 299
OY 317 LTSYIDSKYVPNNNTKLDVEKKTALSSVKNKTIVYEQRPNNRPTANLQSMFTNIDTKNHT 376
Db 300 QVAFAKRKNATTDKTAYKMEVTLGNDYSEELIVDY--GNKKAOPLISTSTNYINNECLS 356
OY 377 VEOTIYIN-PLRYSAKETNV-NISGNGDEGSTIIDSPIIKYKVGDNQNL-----PDSN 429
Db 357 RNMATAYVNOBKNTYTKQTFVYNLT-----GYKFNPNAKNEKIVETDQNOQFVDSFTPDS 411
OY 430 RIYDYSEEDVTNDVYAQLGNNNDVNIENGNIDS--PYIIKYI-----SKYDPNKDYTT 482
Db 412 KIKDVTQDFVYI--YSNDNKTATVDLMKGOTSSNKQYIIQVAYAPDNSSTDNGKIDY-- 466
OY 483 IQQVTVMQTTINEYTGERTASYONTIAFSTSSGOGGDLPEKTYKIGDYVWEDVDKDG 542
Db 467 ----TLDTKTKYSW---SNSYSNVNGSSTAND-----QKKYNLGDYVWEDVTNKDG 511
OY 543 IONTDNEKPLSNVLYTLTYPDGTS-KSVRTDEGKYOPDG 582
Db 512 KQDA--NEKGIKGYVILKDSNGKELDRITTDENGKYOFTG 550
```

Search completed: March 14, 2003, 12:58:23
Job time : 30.0323 secs

[illegible]

```

QY 1 SDEKKNDVYNNOISINDDNNQIIEKTEENNYDGIKRSDEBTESTJTVNDENEATFLOK 60
Db 51 SNEKSNNSSVSAPAKPDITVNV-----SDPKTSNNT--NGEVSVAQ 91
QY 61 TPQONTHL7EEVKKSSSVESSENSSIDTAQOESHHTINEESVOISDNVEDSHVSPFANS 120
Db 92 NP-----AQOETOSSSTNATTEETPVTGEVTTTTTNOANTPATQOS--SNINAEVLV- 143
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTTSQPSGYTNIDEKISNOBELLN-LEINENKA 179
Db 144 --QTSNETTFPNDNTNV-----SVNPSQNSNNAENVSTTQDSTEAFTSN--NES 189
QY 180 RPLSTSAOPSIKRVTVN-----QIAEQGSNVNHLIKVTDQ--SITEGYDSEGV 228
Db 190 APOSTDASNKNKDVAOAVNTVAPSARMAFSLAADAAPAGTITNOLTNTVTVGI--DSGTT 248
QY 229 IKAHNAENLIYDVEFEVDVKVSGDPTMYVDIDKNVPSDLTSPFIPIKDKDSGEIANG 288
Db 249 VYPHQAGVYKLVNGSVSPASAVKGTFTKITYKEKELNNGVSTAKVPPIMAGD-QVLANG 307
QY 289 TYDNKKOITFTFTYDVYKYEENIKAHLLKTSYDIKSKVYNNNTKL-DVEYKTALES--VN 345
Db 308 VIDS--DGNVYITFTYVNTKDDVKATILMPAID----PENKKRGNVILAIGISGTIAN 362
QY 346 KTIIVGYORPNEENTANLOSMEFTNIDTKNHVEQTIYINPL--RYSAKETNVNISGNDE 403
Db 363 KTVLVDYKYGKFEYMLSTIGQIDKTNNTYRQTIYVNPSPGDNVIAVLGTOKENETDS 422
QY 404 GSTIIDBSTIIKVVYVGDQNLPPDSNRIDYSELEYVNTDDVYAOLGNNNDVINNEG---- 459
Db 423 NALIDQOQTSIKVYKVDNAADLSESYFV--NPENFEDVNTSVNITPPNPQYKVEFMTPPD 481
QY 460 NIDSPYIKIVSKVYOPN--KDDYTTIQOQVYTMOTITINEYGE--FRTASYNDAIESTGS 516
Db 482 QITTYIYVNVGHIDPNKSGD-----LALSTLIGYINSNIITWRSMSWMDNEVAFFNGSG 534
QY 517 QOGG-DLP--PEKTYKIGDY--VWEDVDKD-QIONTNDN 549
Db 535 SGDGIKRVNPEQPDPEGEIPIRPEDSDSDSPSDSGSDS 573

```

RESULT 4

US-08-856-253-7

Sequence 7, Application US/08856253

Patent No. 6288214

GENERAL INFORMATION:

APPLICANT: Hook, Magnus

APPLICANT: Patti, Joseph M.

APPLICANT: House-Pompeo, Karen

APPLICANT: Stahanam, Narayana

APPLICANT: Symersky, Jindrich

TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS

TITLE OF INVENTION: AND METHODS OF USE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: U.S.

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856,253

FILING DATE: Concurrently Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/017,678

FILING DATE: 16-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REFERENCE/DOCKET NUMBER: 33,928

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 345 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-856-253-7

Query Match 11.6%; Score 349; DB 4; Length 345;
Best Local Similarity 30.8%; Pred. No. 1.2e-13;
Matches 102; Conservative 61; Mismatches 136; Indels 32; Gaps 15;

213 VDDQ--STEGYDDSGVYKAHKAENLIYDFEVDVKYKSGDTMIVDIDKNTVPEDLMD 270
25 ITNQLTNVVGI-DSGTYVYPHQAGYVKLNGFSPNSAVKGGTFKITYPKELNLGVTS 83
271 SEFTPIKXNGSEIIATGYDNKNOITTYTDYVDKYEINIKAKLITSIDSKYPPNN 330
84 TAKVPIMAGD-QVLANGYDS-DGNVITFTDYVNTKDDVKATLIMPATID----PENY 137
331 TKL-DVEYTAUSS--VKTTIVEYQRPENRPNATQSMFTNIDTKNHYEQITLYNPL- 386
138 KKGANTLTATIGSTANTVLDYKRYGKFMLSIKGTIDQIDKNTNRYRQTIVYVPSG 197
387 -RYSAKETVNVISNGDEGFTIIDSTIIKRYVGVGNONLPSNRIYDSEYEDVNDYV 445
198 DNVIAVLGNLKPNTDSNALLDQONTSTIKVYVDNAADSESYFV-NPENFEDVNSVN 256
446 AOLGNNDVNING---NIDSPYIKVYSKYDPN-KDYTTIQQTVMQITINEYTG- 499
257 IFPEPNQKVEFNTPDQDTTPYIVVNGHIDPNSKGD-----LALNSTLYGVNSNI 309
500 -FRTASYDNTIAFSTSSGGQG-DLP--PEK 526

DB 310 IIRSMWDEVAFNNSSGSGDGDIDKRVPEQ 340

RESULT 5

US-08-714-402-2

Sequence 2, Application US/08714402

Patent No. 5910441

GENERAL INFORMATION:

APPLICANT: ROCHA, Claudia

APPLICANT: FISCHETTI, Vincent A.

TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING

TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,402

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REFERENCE/DOCKET NUMBER: 39,300

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1112 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-714-402-2

Query Match 8.6%; Score 259.5; DB 2; Length 1112;
Best Local Similarity 21.2%; Pred. No. 9.9e-08;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;

125 SNTSGKEENTIPQPNKVEDSTTSOPS-----GYTIDKISNODEL--NLPIV 173
43 ANQSGFEIKKVDONKPLPGATSLTSKQKGTSVQSFNSKQIVDAQNLPGTYTLK 102
174 E-----YENKARPLSTSAQPSIKRYVYV-----QLAAEGS-NVHLIKVTD--OSTE 220
103 EETAPDGYDTSRTWTYVYENGYTRKLVENPYNGELISKAGSKDVSSLDLENPKMSVVS 162
221 GYDSDS--GVIAHAENLIYDFEVDKVS-----GDYMTVIDDKTVSDTLDSF 272
163 KYKTEVSSGAADFYRNHAAYFMKSEELKOKDKSETINPEDTVLQDRLEKFGISOD- 221
273 TPKR-IKDSNGEIIATGYDNKNOITTYTDYVDKYEINIKAKLITSIDSKYPPNN 331
222 -IKRIIYDSANSLAIGKVAENHQLYFTFDYIAGDKVQLABELSLFLENKEYLEN- 278
332 KLDVEKTLASVKNKTI-----TVEYQRPNE-----NTANLQSMFTNIDTK 373
279 -----TSISNKRSTIGGEITYKGVNVLYGNESTKESNYITNGLSNVGGSIESTNE 331
374 NHTVEDQITNPLRYSAKETVNVISNG-----DEGSTIIDSTIIKRYVGVGN 423
332 TGEFVWYVYVNPRTNIPYATMLMGFGRAKRSNTSLENDANTSSAELGEIYVEPEGE 391

QY 424 NLPDSNRITYSEEDVT-----NDYAOQLGNNDV-----NINGN--IDSPYIIKYSK 472
DB 392 KLPSSYGV-----DYTKLRLTRDITAGLNGFOMTKRORIDGNNINOMAKFIKYGK 444
QY 473 YDPKKDDYTTIQOTVMTQNTINEXTGERFTAS-----YDNTIAFSTSSGGGCG-- 520
DB 445 TDOS-----GKPLVQSNLASFEGASEYAFTPVGNGVYFQNEIALSPKSGSGSKSE 497
QY 521 ----DLPEKTYKIGDYMEDVDKGIQNTNDNEKPLSNLVLTLPDGTSKSVR--TDE 574
DB 498 FTKPSITIVANLKRVAQLRFRKMSIDNV-----PLPEAFELRSSNGNSQKLEASNT 549
QY 575 DGXYOF 580
DB 550 QGEVHF 555

SULT 6
US-09-327-536-2
; Sequence 2, Application US/09327536
; Patent No. 6355477
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; APPLICANT: ROCHA, Claudia
; TITLE OF INVENTION: FIBROECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A
; FILE REFERENCE: 022927-008
; CURRENT APPLICATION NUMBER: US/09/327,536
; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: US 08/714,402
; PRIOR FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: SFBP gene
US-09-327-536-2

Query Match 8.6%; Score 259.5; DB 4; Length 1161;
Best Local Similarity 21.2%; Pred. No. 1e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;

QY 125 SNTSGKEENTIEQPNVKEDSTISOPS-----GYTNIDEKISNODEL--NLPIN 173
DB 43 ANKGSEIKKVDONNNPLPGATSLTSKDGKTSVQSFSTNDKGIYDAQNLQGTITLK 102
QY 174 E-----YENKAPLSTTSAPPSIKRYTVN-----QLAAEQGS--NVNHLIKYTD--OSITE 220
DB 103 EETAPDGYDKTSRTMTVTYVYENGTYTKLVENPYNGEIIISKAGSKVSSLSOLENFKMSYVS 162
QY 221 GYDSDS---GVYKAHDAENLIDYVTFEYDVKVS-----GDTMTVDIDKNTVPDLDSEF 272
DB 163 KYGTETVSSGADYRHHAAAFKMSFELKQDKSETINPGDTFYQLDRALNPGLISGD- 221
QY 273 TIFR--IKONGSEIIATGYDNKNOITTYTFTDYVDKYENIKAKLTSYIDSKSVPNNT 331
DB 222 -IPKIIYDASNPILAIGKYHAENHQLIYTFDYIAGLDKQLSAELSLFLENKEVLEN- 278
QY 332 KLDEYETALLSVNKTI-----TVEYORPNE-----NRRANLOSMFTNIDTK 373
DB 279 -----TSSINERSTJGGOEITTKGVNVLVGNESKESNYITNGLSNMGSIESTYNT 331
QY 374 NHTVEQTIYINPLRYSAKETVNIISNG-----DEGSTIIDSTIIKVKYKGDNO 423
DB 332 TGEVWYVYVNPNTNIPYATNMLMGFRASNTSDLENAMTSALGELQVYVEPEGE 391
QY 424 NLPDSNRITYSEEDVT-----NDYAOQLGNNDV-----NINGN--IDSPYIIKYSK 472
DB 392 KLPSSYGV-----DYTKLRLTRDITAGLNGFOMTKRORIDGNNINOMAKFIKYGK 444
QY 473 YDPKKDDYTTIQOTVMTQNTINEXTGERFTAS-----YDNTIAFSTSSGGGCG-- 520

DB 445 TDOS-----GKPLVQSNLASFEGASEYAFTPVGNGVYFQNEIALSPKSGSGSKSE 497
QY 521 ----DLPEKTYKIGDYMEDVDKGIQNTNDNEKPLSNLVLTLPDGTSKSVR--TDE 574
DB 498 FTKPSITIVANLKRVAQLRFRKMSIDNV-----PLPEAFELRSSNGNSQKLEASNT 549
QY 575 DGXYOF 580
DB 550 QGEVHF 555

RESULT 7
US-09-134-001C-5080
; Sequence 5080, Application US/091344001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5080
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080

Query Match 7.3%; Score 220.5; DB 4; Length 3696;
Best Local Similarity 20.3%; Pred. No. 9e-05;
Matches 134; Conservative 106; Mismatches 241; Indels 180; Gaps 31;

QY 2 SDEKNDVINNNOSINTDNNQIIEKERTNNYDIEKRSERDTESTYIVDENEATFLQKT 61
DB 718 ADQCKT-LIRNNHATTEEONEAIRQYEAHSSDAIAKGEAFDTDTVTEARNDNGKLIAT 776
QY 62 PODTHLTREEKSSSVESNSSSIDTAAQPSHTTN--REESVOTSDHVEDSHVSDFANS 120
DB 777 DVNP--TKKARAAYATNSANSKIKIDINNTQATLDERDAIALVNSKD-----E 826
QY 121 KIKSNTSGKEENTIEQPNVKEDSTISOPSGYTNIDEKISNODELINLPINEYENKAR 180
DB 827 AIONINTRAQG-----NDVYTEAQNNG--TNTIQOV----- 854
QY 181 PLSTTSAPPSIKRYTVNOLAEOGSNVNHLIKYTDOSITEGYDSDSEGVYKAHDAENLIDY 240
DB 855 PLTVKKNQNAL--ATINAKADEQ-----KRLQANNMATE-----EKADAEKRYNE 899
QY 241 VTFEVDKVSQGTMTVDIDK--NTVPDLDTSFT--IPKIKONS-----GELIATGYDNK 293
DB 900 AVITANONITNA--TTNRDVDAQTTGSGIISAIIPATRIKEDARAAYAKAIAQNOQINS 958
QY 294 NKQITVFTDVVDYENKKAH-----LKLTSYIDSK-----VNN 329
DB 959 NNMA--TTEEKEDALNOVEAHKQALAIATINOAGSTQOVSSEAKNNGNINTINODQPAVAKN 1016
QY 330 NTKLDVEYK-----TALSSVKNKTI--VEYORPNEPNTANLOSM 366
DB 1017 NTKLILQKNGEKKSAIAQTPDATTEKEQEAHSVSAVQTGITHINANEN----- 1068
QY 367 FTNIDTNNHTVEQTIYINPLRYSAKETVNIISG-----NDEGSTI 407
DB 1069 -DVIDOELSNAEQIIT-----THTVNVQKRPQARQALIAKTNERKOSAINSDNEGTI 1118
QY 408 IDDSITIIKVKYVGNQNLPSNRITYSEEDVTYNDYAOQLGNNDVNIIN-----GN 460
DB 1119 EEKOKAIO--SLNDAKNLAD--EQITQASQNVNVALNIGISNISKIQTNETKQOARDQ 1175

QY 461 IDSPILIKVI-----SKYOPNKDDYTTIOQVTMTTINEIGEPTASTYDNTIARST 513
 DB 1176 VNOFOEKEAPLNTSPHATQOEKODALTRLTQ--AKETALNDINOAGTNOANDTALTSGI 1233
 QY 514 SSGOG-QODLPEPTYK--IGDYWEDVDKXGIONTND--NEKLSNVLVTLYTPDGT 567
 DB 1234 QNIONTOYVNRKROEAKTTINDIYQO--HKOSIONNDATTEKEEVANLVNASOONVIS 1291
 QY 568 K 568
 DB 1292 K 1292

RESULT 8
 US-08-973-462-9
 ; Sequence 9, Application US/08973462B
 ; Patent No. 6191270
 ; GENERAL INFORMATION:
 ; APPLICANT: DROTHLE, PIERRE
 ; APPLICANT: DAUBERSIES, PIERRE
 ; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
 ; FILE REFERENCE: 0660-0125-0 PCT
 ; CURRENT APPLICATION NUMBER: US/08/973.462B
 ; CURRENT FILING DATE: 1998-02-06
 ; EARLIER APPLICATION NUMBER: PCT/FR96/00894
 ; EARLIER FILING DATE: 1996-06-12
 ; EARLIER APPLICATION NUMBER: FR 95/07007
 ; EARLIER FILING DATE: 1995-06-13
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 630
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
 US-08-973-462-9

Query Match 6.9%; Score 209; DB 4; Length 630;

Best Local Similarity 19.5%; Pred. No. 4.7e-05;
 Matches 116; Conservative 121; Mismatches 206; Indels 152; Gaps 28;

QY 2 SDEKNDVNN-----NOSINTDNNQIKKEETNNYDGIKRSRDRST- 47
 DB 59 SDELFELLNSVDVNGEVENILEESQVNDIDFNSLVKVCQEQHNVKEKVESEEND 118
 QY 48 -----TNDENE-----ATFLQK-----TPQONTHTLJTEEVK 74
 DB 119 EESVEENVEENVEENDGSAVSSVESIASVDESIDSIEENVAPLVEIYAPVEIY 178
 QY 75 ESSVSSSSNSIDTAOPSHTTINRESVOTSDNVEDS--HVSDFANSKIKES--NTE 128
 DB 179 ABSVSVASVAPSVESVAPS-----VEESV--AEVNEESVAENVEIYAPSVESV 231
 QY 129 SGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKIS--NODLNLPLINE-----YENKARPL 182
 DB 232 ESAVNEESVAENVEESVAE-----NVEESVAENVEEYIAPLVEESVAPLVEIYAPT 285
 QY 183 STTSAPSTIKRVTVNOLAAEOGSNNHLIKVTQSTIEGIDDESGVYKADAENLIYDVT 242
 DB 286 VESVAPLVEIYAPSVESVAPSV-----EESVAENVEESV-----AEVNEESVA 331
 QY 243 FEVDKVKSG--DTMTVDIKNTVPSDLTDSFTPIKIKNGEIIATGTYNKKNKOI--- 297
 DB 332 ENVEESVAENVEESVAENVEEYIAPSV--VEEYIAPLVEESVAENVAENLNDNLNLGG 389
 QY 298 --TYTFTDYV--DREYENIKAHKLKLTSTYIDKSKVPNNNTKLDVEYKTLASV----- 344
 DB 390 IETFEIKDSILNEIEVEKENV--VTTLIEKVE-----ETLAEVSTFTSNIIEIEIOENTIT 442
 QY 345 NKTIYEQPNNENRNTANLOSMFTNIDTKNHTVEQTIYINLKRSAKETVNNVNSIGNDEG 404

DB 443 NDTIEKLEELHEN---VLSALENTQSEEEKREVIDYIEEYKEEVAATLLETVEQAESE 499
 QY 405 STIIDSTIIKYY-----KVGDN-QNLPDS--NRKYDSEY-----EDVYND 443
 DB 500 S-----ESTITTEFLENNAVESNEKVAENLEKLETQFNVLVDVEETVEISGLENN 555
 QY 444 D-----VQOLGNN-----NDVINFGNIDSPYITIKVISKYDPNKDDYTTIOQTV 487
 DB 556 EMDKAFSEIFDPNVKGIQENLLTGMRFSIETSIIVIOSEKVLNENNVSSIIDNI 610

RESULT 9
 US-09-134-001C-4318
 ; Sequence 4318, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134.001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 4318
 ; LENGTH: 676
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4318

Query Match 6.7%; Score 202; DB 4; Length 676;

Best Local Similarity 22.1%; Pred. No. 0.00013;
 Matches 116; Conservative 82; Mismatches 191; Indels 136; Gaps 23;

QY 2 SDEKNDVNNNOSINTDNNQIKKEETNNYDGIKRSRDRSTTNDENATFLQKT 61
 DB 77 SPEERLDKVDNOSIDRIIND-----AKDKNNH--LKSTSSAATSSITTEDDT-----S 123
 QY 62 POONTHTLEEEVSSSVSSNSIDTAOPSHTTINRESVOTSDNVEDSHVSDFANSK 121
 DB 124 EKNDNMTKDLKILDLDSIAKNVDNRQOEGSASKPSDS--TTIEKDS-----NNK 175
 QY 122 IKESN-----TESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNODLNLPLINE 175
 DB 176 VHDTNASTRNATTDSDSESVYDKLKIQDFKSDS-----NNKLSQSDQOASPSKN 228
 QY 176 ENKARPLSTISAPST-----KRYTVNOLAAE-----OGSNVNLKIKYTD-----Q 216
 DB 229 ENKKEESSTTNNOSDSKDKSNDGRRSLERIASDTDOIRSKDOH---VTDEKODIO 285
 QY 217 SIT---EGYDSEGVYKAHDAENLIYDVT-----EVDDKVKSGDTMTVD----- 258
 DB 286 AITRSLOGSKIEKALAKVSDNQPDSNVIINKMLNLSRLDKVEDNNTLSDDKQALK 345
 QY 259 --IDKNTVPSDLTDSFTPIKIKNGEIIAT-----GTYDNKNO 296
 DB 346 OEIDKTKOSIDROENIILIDOLNGASNNKQAKTEIDILNSVFSKNEVEDIMKRIKTGRSNE 405
 QY 297 ITTFTDYVQKYEIKAHKLKLTSTYIDSKYPNN-----NTFLDVEYKTLASV-- 343
 DB 406 IANOIAKOIDGLALTSSDDILKSLMDSKOKESILIKOLLTRGLANDADRIAKLLSONL 465
 QY 344 VNKTIYEQRP--NENRNTANLOSMTNII--DTKNHTVEQTIYINLPRYSAKETVNNVNSIGN 400
 DB 466 SNSQIYOLKRRHFNFSQGTATADDLINGVINDAKD-----KROAIEITLQTRIN 513
 QY 401 GDEGSTIIDSTIIKYYKVDGNQNLPSNRITYD--VEEYEDVYND 443
 DB 514 KDKAKIAD--VIARVOK-----DKSDIMDLHSAIEGKAND 548

RESULT 10
US-09-134-001C-5106
Sequence 5106, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIORITY APPLICATION NUMBER: US 60/064,964
PRIORITY FILING DATE: 1997-11-08
PRIORITY APPLICATION NUMBER: US 60/055,779
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5106
LENGTH: 1177
TYPE: PR1
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5106

Query Match 6.7%; Score 201.5; DB 4; Length 1177;
Best Local Similarity 19.1%; Pred. No. 0.00029;
Matches 131; Conservative 121; Mismatches 242; Indels 191; Gaps 30;

QY 7 NDVNNNSITDDNNQIKKEETNNYDGIKRESEDTSTTNDENATFLQKT----- 61
DB 134 NNYNNNSVFNQNNKSSQQRKSTOSENNKANKTKNTSTS---PEFTYLNHSFKSE 189
QY 62 -----PDQNHLPTEEVKESSESVSSSIDTAQOPSHHTINREES 102
DB 190 VPSAIFGTRKRPLENGVYIPEHKELENDKEIVQDEYSHSTKSDASKVNSNDNNIEK 249
QY 103 VQ-----TSDVVEDSHVDFANSKIKESNTESGKEENTY----- 136
DB 250 NQAKKQOQTAQTESSEMMHVEKSNVQTKRKTPN--YSKVDNTINENIYASQIVEEI 307
QY 137 -----EOPNKVYEDST-----TSOPSGYT--NID-EKIS 162
DB 308 RREBERKVLQRRFKKALQOKRQONQOSEDSIOKAIIDEMAKQAHYTGESSIDLENES 367
QY 163 NQDELNLPINEYENKARPLSTTSNOPSIRKVTYNOGLAEQSGNNHILIKYTDOSITG 222
DB 368 NQDSSSN--SLEKQSSNNIDNKEAQNTPLENEEIDLDITSDV--YKVNTEE--TESK 421
QY 223 DDESGVIAH-----DAENLIYDVF--EVDKVKSGDTMTVD--IDKNTVPSDLTDSFTI 274
DB 422 NDEDLVSSNHTHSNDAAE--VEDAEYHLLDNRQONQNSODDITSSKSSSTSMIDNAIS 479
QY 275 PKIDNGSEIATGTYNKKNQITV-----TFDYDYKYENIKAHKLTL 318
DB 480 ASVNNMTER--AKSNEDKNQTEITHLDGTSKAKVSEKIESENHNLHODKNVK--LKNV 535
QY 319 SYIDSKVPPNNNTK-----LDVEYKTAASSVVKITVEYQRENNENTALQ 364
DB 536 NSLWSSNDGTQTRQRFQGGSRPENVLTPSDKRMQONKRVKVSPELKEKQANNAHR 595
QY 365 SMFTNIDKNTVEQTIYINPLKYSKAKETNPN--ISNGDESGTITDSTIIKVVYVGNQ 423
DB 596 K-----DESKKSEEFKIN-----TNRETDSNSYESNGIEND--MNSSDEHYVITPQK 644
QY 424 NLPDSNRILYDSEYEDVTNDYDQALGNNDVNNINEGNIDSPYLIKVISKYPNNKDYTTI 483
DB 645 ODEQLOKQDQDFHEF--NANHAKINNSNET--GN-----QENI 678
QY 484 QQTVMQITINEYGEFTATSYDNTIAFSTSSGQGGDLPEPKTKIDY-----VMEVD 539
DB 679 SHSKRSQYSTRE-----SKNIDTQTSNSTSNQFORIRKGPNIKLPSTOLLEAPHEP 732
QY 540 KDGIQNTNDNEK--PLSNVLVTLTP 563

DB 733 KD--QDWIDNKKQELNDALYYFNVP 755

RESULT 11
US-09-268-347-49
Sequence 49, Application US/09268347
Patent No. 635182
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT APPLICATION NUMBER: US/09/268,347
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 49
LENGTH: 2314
TYPE: PR1
ORGANISM: Haemophilus influenzae
US-09-268-347-49

Query Match 6.5%; Score 197; DB 4; Length 2314;
Best Local Similarity 21.1%; Pred. No. 0.0012;
Matches 135; Conservative 101; Mismatches 261; Indels 144; Gaps 30;

QY 4 EEKNDVNNNSIN--TDNNQIKKEETNNYDGIKRESEDTSTTNDENATFLQKT 62
DB 379 ETOADKLTDNNNIGVNTNNGKLKYLAKNLGLSGL-----TVSTKNTLASE---KVTY 428
QY 63 QDNHLPTEEVKESSESVSSSIDTAQOPSHHTINREESVQTSN---VEDSHVSDFA 118
DB 429 GSGNNTALQGGGLFTFTTNASTD-----KTYVGDGLKLTGNTSNTALEDT--TRAT 479
QY 119 NSKIKESNTESGKEENTIEQ---NKVEDSTSPSGYTNIDKISNDELNLPINE 174
DB 480 KDKIGFSKAGTVDEN---KPYLDKDKLVGNSLTNNGGLT--VNNITGSKNKOIYV--- 531
QY 175 YENKARPLSTTSNOPSIRKVTYNOGLAEQSGVNNHILIKYTDOSITE---GYDSSGVIAK 231
DB 532 -----GADGIFADVN-----VWVSAAFEGTTRITEEIGFADAGKVDK 572
QY 232 H-----DAENLI--YDVFVVDKVKSGDTMTVIDKNTVPSDLTDSFTIPKIR---DNSG 282
DB 573 KSPYIDKQLOVGVYKTKRDSGINAGDOKISNVKAT---DQDAVYTKQKQVOODADG 629
QY 283 EIAATGYDNKKNQITV-----YTFDYDYKYENI---KAHLKLTYSIDSKVPPNNNTKLD 334
DB 630 ALQSFIDEKQGEFTISNLYSNGNTPMTFETITFAGEGKISINDIAKGV-----KVG 684
QY 335 VEKTAALSSVVKITVEYQRENNENTALQSFNIDKRNHVEQTIYINPLKYSKAKETN 394
DB 665 IDPGLTTPKLTV-----GSDKDGKTLVIEQVASGNGKNTIRGVSPLPSI--TN 735
QY 395 VNISNGDESGTITDSTIIKVVYKVGDNQNT-----PDSNRILYDSEYEDVTNDYDQALQ 449
DB 736 AGCVATTQGNITTSDEKSKAASIGDILNTGFNLKNNNSNGFVSTYTVTFID-----G 791
QY 450 NNDVNNINEGNIDSEYIYIKVISKYPNNKDYTT-----IQQVTVMQITINEY 496
DB 792 NATKAKVYDEFNQSKV---TYDVANDEKTIETLGDNGKTKNKIGVKTTLTTNNANGK 847
QY 497 TGEFTATYD-----NTIAFSTSSGQGGDLPEPKTYAI--GDYVED-----VD 539
DB 848 ATNFTNDNALVNAKDAIENLNTLAKELHTTKGRADALQ--TFKVKKDGATVDETITVG 906
QY 540 KDGIQNTNDNEKPLSNVLVTLTPDGTSSKSVRTDDEGKYOF 580
DB 907 KDGTQNT-----GKTYNTLTKGKENGTLVATINNDGYTVF 939

RESULT 12
US-09-134-001C-3716
Sequence 3716, Application US/09134001C

Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3716
LENGTH: 1335
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3716

Query Match 6.5%; Score 195; DB 4; Length 1335;
Best Local Similarity 20.8%; Pred. No. 0.00081;
Matches 160; Conservative 97; Mismatches 254; Indels 260; Gaps 43;

7 NDVYNN--NOSINDDNO-----IK--KEETNNYDGIKESDRTSTVDENEATF 57
100 NEISSOKOQSLSTDANONOTSVTKNOQETN-----DLTQCKSTDTNQLDETQSV 154
58 LKTPDQD---NTHLEEEKSSSVSSNSIDTAQOPSHTTINREESVQTSNDVEDSHV 114
155 AKENKDKGAMANNQOQKMTASOPSENQAIET-----QTSNDNESQOKSOQV-TSQ 208
115 SDFANSKIKESNT-----ESGKEBNTIEQPNKYKEDSTTSPSGY----- 154
209 NETAPKSTNWSAGNFYDDEDDSDTDLEPISLNNVNTSKQTSYKKEPAQRYT 268
155 TNI--DEKISNDELINLPINYEKKARPLSTTSAPSKRTYVNOALAEQSNVHLK 212
269 TMTVKETASNOATIDTQTFPSATQPRIVYSV-SQKTSLEPYPYKRVSSINNYR 327
213 VTQDS---ITEGYD-----SEGYIKAHDAENLYDVTFEVDKVKSGDTWT 256
328 KKNMKAPEEEDYTSYFPRKYGRNGVGREGLV-VHDTAN-----DNST 370
257 VD-----IDKN-----TVPSDLT-----DS 271
371 IDGEIAFMKRNNTAFVHAFVDGNRIETAPDYLWSGAGYGNQRFINVEIYHTHDYS 430
272 FTIPRIKNSGEIAT-----GTYNKKKQITYTF-----TDYDKYENIKA 313
431 FA--RSMNNYADYAATQLOYNLKPDSEADGQVWTAHLSNLTGTDHADPHOYLS 488
314 H-----LKITSYIDSK--VP--NNNTKLDEYKKTALSSVNTTIVEYOR----- 354
489 HNYSAELDYLDIYEKTLITKQVAPMGSTSTKPSQSPKPS--GGTNKKLTVSARGAQK 547
355 PNEENTANLOSMTNI-DTKNH---TVEQTYI-----NPKRIYAKETN----- 354
548 PTNN-----GLYTYVDDGKHGTQVOKTSLVTKTATLGNKFYLVEVDSNKKYGVW 601
395 -----VNISNGDEGSTI-----IDSTIIKVKYVDNOMLPSNRLYD 434
602 QGDVYVNTKAPYKQVOTNVKAGSTLYVPKCTPKQVASKVSGTG-----NOTFKA 653
435 SEVEDYNDYDQAGNNNDVNFNIDSPYIIKIVISKYDPNKDDYTTIOQVYTMOT--- 491
654 TQOQIDKATLYL---GVNKGSGWISKYLLTATASKPSPTKPS--TNNQLTYTNNSGVA 708
492 TINETYGERRTASD-----NTI--AFSTSSGQGGDLPEKTYTIGD-----YWED 537
709 QINAKNSGLYTYVDTKGTQNTQIORTLSVTKAATLGD---KKFYLVGDNVTGTNGYV 765
538 VPKDQONTNDNKKPLSNVLYLTLY-----PDGSKSVRFDEDEK 577

Db 766 QD-EVIYNT-----AKSPKINQTYNVKRGVKLHTVPMTYNOVAGTVSGK 810

RESULT 13
US-09-134-001C-3159
Sequence 3159, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3159
LENGTH: 10182
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 6.5%; Score 195; DB 4; Length 10182;
Best Local Similarity 19.8%; Pred. No. 0.01;
Matches 133; Conservative 102; Mismatches 257; Indels 180; Gaps 26;

2 SDEKNNDVYNNOSINTDNNOLIKKEETNNYDGIKESDRTSTVDENEATFLOKT 61
5081 SONE--SVTNESYVINAEPKQHAFTALNNAKEIYNEQATILDANSOKAAILTTN 5138
62 PODNTHLEEEKSSSVSSNSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANSK 121
5139 ALDG---EQQLRAK--ENAOEITLNLQTDQNSKGLINSQTFREVASOLAKK 5192
122 IKESNTSGKEENTIEQPNKYKEDSTTSPSGYTNIDEKISNDELINLPIN--YEN- 177
5193 -----ELKKVME-----OLNHLINKNOMINSKRFINDANOQAYSNA 5231
178 --KARPLSTTSAPSKRTYVNOALAEQSNVHLKYVDQSTIEGYDSEGYIKAHDAE 235
5232 IASAEALKKNSQNPBLDKQVTEBOAINNINSAINL--NGEAKLTAKKEDVAST----- 5283
236 NLIYDTFEVDKVKSGDTWTVDIDKNTVPSDLTDS-----FTIPRIKNSGEIATG 288
5284 NNLSGLTNE--QTKENQAVNGAQTRDOYANKRLDAEALDQSMQTLRDLVNNQNAIHS 5341
289 TYDNKKKQITYTFDYDKENIKAHKLSTYIDSKSVPNNTKLDEYKTL--SSVK 346
5342 NFENEDSTOKNTDNDID-----NGSYTYTGQHNPELKNKSTTIDQ 5380
347 TIVERQRPN--ENRTANLOSMTNIDTKHTEQTYIINPLYSKAKETVNTISG----- 399
5381 TIS-----RINTAKNDLHGEVKLORDGTANOELGQLYNDPKSGSESLVNSNTRSEV 5436
400 -----NGDEGSTIIDSTIIKVKYVDNOMLPSNRI--- 431
5437 EEHLNFAKSLNNAMKOLRDKVAEKTNVKQSSDYINDST---EHQRGVQALQEAENLINE 5493
432 -----YDSEYE---DYRTDDYDQALG-----NNNDVYN-----FGNI 461
5494 IGNPITLNSKIEOKLOOLDTDAONALGSHLEAKKNATITGINKLALANDAQOKAIENY 5553
462 DSPYIIKIVISKYDPNKDDYTTIOQV-----VTMOQTYNETYGERP-----TASYDNTI 512
5554 QAOQOTIPAVNOQLTLDBREITAMQALBRKYGQGNVHQSNYTNEDQPKHNTDN-----S 5609
513 TSSGQGGDLPEKTYTIGDYVEDVDK--DGIONTNDNKKPLSNVLYLTLYPDGTSKV 570
5610 VQAGQ-----TIIDKLQDPINNKKEIHOATINQITQTALSGEKLH 5651

OY 571 RTDEGKYQFDG 582
DB 5652 TPOESTNRQIEG 5663

RESULT 14
US-08-973-462-8
Sequence 8, Application US/08973462B
Patent No. 6191270

GENERAL INFORMATION:

APPLICANT: DROUILHE, PIERRE

APPLICANT: DAUBERSIES, PIERRE

TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

FILE REFERENCE: 0660-0125-0 PCT

CURRENT APPLICATION NUMBER: US/08/973.462B

EARLIER FILING DATE: 1998-02-06

EARLIER APPLICATION NUMBER: PCT/FR96/00894

EARLIER FILING DATE: 1996-06-12

EARLIER APPLICATION NUMBER: FR 95/07007

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 8

LENGTH: 1786

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Polypeptide

US-08-973-462-8

Query Match 6.4%; Score 192.5; DB 4; Length 1786;
Best Local Similarity 18.9%; Pred. No. 0.0016;
Matches 128; Conservative 138; Mismatches 261; Indels 149; Gaps 26;
OY 14 OSINTDNNQIKKEETNNYDGIERSDRTSTNVDENEATFLOKTPQDNTHTLVEEV 73
DB 549 ESVEENVEESVAENVEESVAENVEESVAENVEE--SVANVEEIVAPVEEIVAPVEEI 606
OY 74 KSSSVESNSSIDTAQPSHTTIRRESVQTSNVEDSHVSDFPNSKIKESNTESGKE- 132
DB 607 VAPSVESVAPSVESVVE--NVEESV--AENVEES--VAENVEESVAENVEEI 654
OY 133 -----ENTIEO--PNKYKEDSTSPSGYTNIDKIS--NODELINLPINYEKA 179
DB 655 VAPVEEIVAPVEEIVAPSVESVAPSVESVVEENVEESVAENVEE--SVANVEESVA 712
OY 180 RPLSTSAO-----PSIKRV--TVNQLAEOGSNNHLIKYVDQSTTEGIDSEGYI 229
DB 713 ENVEESVAENVEEIVAPVEEIVAPVEEIVAP--SVESVAPSVESVVEENVEESV-- 767
OY 230 KAHDAENLIYDVEYDQKVGSG--DTMTVIDKNTVPSDLTSTTIKIKNSGEIAT 287
DB 768 ---AENVEESVAENVEESVAENVEESVAPVEEIVAPS--VEESVAPSVESVVAENVAT 821
OY 288 GYDKNKNOI-----TYFTDYV--DKYENIKAKLKYTSIDKSKYNNNTKLDVYKTA 341
DB 822 NISDULSNLLOGITEETEEKSIDLNEIEVKEVNTT-----ILENVEETTAESVTF 874
OY 342 SSV-----NKTIVEXORPEN--RTANLQSMFTNIDTKN----- 374
DB 875 SNILEIQUENTITNDTIEKLELHENVLSALENTQSEEEKKEVIDYIEVKEEVATTL 934
OY 375 -HTVQO-----TIINPLRYSAKETNNVNSGNGDEGSTITIDSTTIKIKYKGDNON 424
DB 935 IETVQAEKSNITTEIFENLEENAVESNENVAENLEKNETVNTVADKVEETVEISG 994
OY 425 LPSDRIVDYSYEDVTNDVYQNLGNNNOVNNFGNIDSPYIIKYSKYDPKADYIT 483
DB 995 ESELENNDKAKFSEIF--DNVKGIOENLIGMFHSIETSIYISSEKVDLENVVSIL 1052
OY 484 -----OQTVMTQTTINTEYGEFTASVDNTIAFSTSG----- 516
DB 1053 DNIEMKGLINKLENISSTBEVQETVTEHVEGQNNYVVDVPAKMDQFLGILNEAGLKE 1112

OY 517 -----QGGDLPEPEYTKIGDYWEDVUKDGIQNTDNKEPLSNVL-----VITL 561
DB 1113 MEFNLEDVFKSSSDYITVEEIK--DEPVQKEVEKTVSIIIEEMENIVDLSEEEKEDLTK 1171
OY 562 YPDGTSKSVRTDEGK 577
DB 1172 MIDAVEESIEISSDSK 1187

RESULT 15
US-08-621-944A-4
Sequence 4, Application US/08621944A
Patent No. 6440425

GENERAL INFORMATION:

APPLICANT: SASAKI, Ken

APPLICANT: HARKNESS, Robin E.

APPLICANT: LOOSMORE, Sheena M.

APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER

MEMBRANE PROTEIN OF MORAXELLA

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Slim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/621.944A

FILING DATE: 26-MAR-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/478,370

FILING DATE: 07-JUN-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-587

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO. 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1833 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-621-944A-4

Query Match 6.3%; Score 190.5; DB 4; Length 1833;
Best Local Similarity 19.7%; Pred. No. 0.0022;
Matches 135; Conservative 114; Mismatches 254; Indels 181; Gaps 34;
OY 3 DEEKDVYNNOSINT-----DNNQIIEKETNNYDGIERSDRTSTNVDENEAT- 56
DB 602 DKDSNMAISINDILNTGKLNKNNNPI--DEVSTYDIYDFANGNATATYHTDANKTS 658
OY 57 -FLOKTPQDNT--HLTEEVKSSSVESNSSIDTAQPSHTTIRRESVQTSN-VE 110
DB 659 KYVIDVAVDDTTHLTGDDNKKLGKTKTKLNKTSANGNTATNEN--VNSSDDALVN 714
OY 111 DSHVSDPNSKIKESNTSGKEENTIE--QPKRVKVEDSTTSQPSGYTNIDEKISNQ----- 164
DB 715 AKDIAENMLAKELHTTKGTADTALQTFYAKKVDENNADANAIYQKKANNGVMTL 774

QY 165 ----DELINLPINEXENKARLSTTSAP-----SIKRYVN--OLA----- 202
 Db 775 TLKGENGLNITDKNNGVTGINTSGIKAGKSTLNDGGLSIKMPGSEQIOVGADGVKF 834
 QY 203 -----OGSNVHLIKVTDOSTEGYDSEGYI----- 229
 Db 835 AKVNNNGVVGAGIDGTRITRIDEI--GFTGNGSLDKSPHLSKDGINAGKKITNIOG 892
 QY 230 ---KAHDA--ENLIYDVFEDDKVS-GDTMTVIDDKNTVPSDLNDSFTIPKIKONG 282
 Db 893 FIAONSHDAVYGGKIVYDLKTELENKISSTAKAONSLHEFSVADGONNFTVSN----- 946
 QY 283 EITATGYDNKKNKOITTFYTVYDKYENITAKLITSYIDS--KVPNNNTK-IDVEYKT 339
 Db 947 ---PYSSYDTSKSDVITFFAG-----EN-----GITTKVKNKGVYRVGIDQTKGLTTPKLT 993
 QY 340 ALSSVNTTITVEYORPENRANLOSMTNIDTKNHTVEOTIYINPLRYSAKETVNIISG 399
 Db 994 VGNNGKGIYIDSON-GONTITGLSNTLANVTNDKGSVRT----- 1033
 QY 400 NGDEGSTIIDD-----STIIKVKVGNONLPDSNRIVDY-SEXYDYNDDYAOQGNND 453
 Db 1034 --EONGNIIKDEDKTRASIVYLSAG--FNLOGNEAVDFVSTYDTVNFAD---GNATT 1085
 QY 454 VNINFGNIDSPYIIKIVISKYDPNKDDYTIQ-----QYVIMQT-----INEX 496
 Db 1086 AKVYTD--DTSKTSKV--YDVNVDD--TTIEVKDKKLGVKTTLTSTGTGANKFALSNOA 1140
 QY 497 TGEFRTASVD-----NTIAFSTSSGOGGDLPEKTYKIGD--YVMEVDVKDGIQNTND 548
 Db 1141 TGDALVKASDIYVAHLNLTSGDIOTPAKGASQANNSAGYVADGKNKYIYDSTDKKIYQAKND 1200
 QY 549 -----NEKPLSNVLVTLTYPDGT 566
 Db 1201 GTVDKTEKREAVAKDLVAQAQTPDGT 1224

Search completed: March 14, 2003, 13:00:50
 Job time : 24.3871 secs

GenCore version 5.1.4-p5-A578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: March 14, 2003, 13:00:30 ; Search time 8.60176 Seconds
(without alignments)
3086.325 Million cell updates/sec

Title: US-09-147-405B-13

Perfect score: 3013 I SSDERKNDVNNNOSINTDD.....PDGTSKSVTDEDKYQFDG 582

Sequence: 1 SSDERKNDVNNNOSINTDD.....PDGTSKSVTDEDKYQFDG 582

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCCT_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	576	19.1	1349	10	US-09-815-242-5898
2	576	19.1	1349	10	US-09-815-242-13137
3	499.5	16.6	932	10	US-09-815-242-5578
4	499.5	16.6	932	10	US-09-815-242-12438
5	455	15.1	841	10	US-09-815-242-5779
6	455	15.1	841	10	US-09-815-242-1751
7	449	14.9	1021	10	US-09-815-242-5471
8	449	14.9	1021	10	US-09-815-242-12544
9	376	12.5	767	10	US-09-815-242-5899
10	376	12.5	767	10	US-09-815-242-13140
11	349	11.6	345	10	US-09-815-820-7
12	319.5	10.6	978	10	US-09-815-242-5456
13	319.5	10.6	1001	10	US-09-815-242-12686
14	308.5	10.2	1018	10	US-09-815-242-5797
15	308.5	10.2	1018	10	US-09-815-242-12838
16	209	6.9	630	9	US-09-742-096-5
17	206.5	6.9	2368	10	US-09-815-242-5635
18	206.5	6.9	2368	10	US-09-815-242-13389
19	201	6.7	2478	10	US-09-815-242-5816

ALIGNMENTS

20	201	6.7	2478	10	US-09-815-242-12967	Sequence 12967, A
21	199	6.6	665	9	US-09-820-843A-107	Sequence 107, App
22	198.5	6.6	807	9	US-09-820-843A-108	Sequence 108, App
23	196.5	6.5	1139	9	US-09-820-843A-15	Sequence 15, App1
24	194	6.4	2150	9	US-10-135-322-17	Sequence 17, App1
25	192.5	6.4	1786	9	US-09-742-096-3	Sequence 3, App1
26	188.5	6.3	1435	9	US-10-153-273-4	Sequence 4, App1
27	187.5	6.2	2344	10	US-09-815-242-12713	Sequence 12713, A
28	186.5	6.2	1111	10	US-09-815-242-12955	Sequence 12955, A
29	186	6.2	836	9	US-09-820-843A-109	Sequence 109, App1
30	186	6.2	861	9	US-09-858-525A-2	Sequence 2, App1
31	186	6.2	871	9	US-09-858-525A-2	Sequence 2, App1
32	182.5	6.1	1421	10	US-09-924-154-13	Sequence 13, App1
33	180	6.0	789	9	US-09-995-587A-1	Sequence 1, App1
34	175.5	5.8	1183	9	US-09-870-759-45	Sequence 45, App1
35	175	5.8	1338	10	US-09-402-100-4	Sequence 4, App1
36	174.5	5.8	2437	10	US-09-815-242-5834	Sequence 5834, Ap
37	174.5	5.8	6281	10	US-09-815-242-12996	Sequence 12996, A
38	173.5	5.8	837	10	US-09-815-242-5883	Sequence 5883, Ap
39	173.5	5.8	875	10	US-09-815-242-13080	Sequence 13080, A
40	173.5	5.8	2122	9	US-09-813-214A-9	Sequence 9, App1
41	173	5.7	2076	10	US-09-815-242-5815	Sequence 5815, Ap
42	173	5.7	2186	10	US-09-815-242-12913	Sequence 12913, A
43	171	5.7	2025	10	US-09-815-242-5703	Sequence 5703, Ap
44	171	5.7	3158	10	US-09-815-242-12611	Sequence 12611, A
45	171	5.7	5795	10	US-09-815-242-12610	Sequence 12610, A

RESULT 1

US-09-815-242-5898

Sequence 5898, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815, 242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191, 078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206, 848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207, 727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242, 578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253, 625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257, 931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269, 308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 5898

LENGTH: 1349

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-5898

Query Match 19.1%; Score 576; DB 10; Length 1349;

Best Local Similarity 30.1%; Pred. No. 9, 8e-24;

Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

```

OY 1 SDEEKNV---INNOSINTDNNQIIEETNNVDIEKRSDETESTVND-NEAT 56
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 55 STNKLNETTASDNOSSDKVDKMOQLNEDNTKNDKEMVSSOGNETTSGNKILIEE 114
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 57 FLQKTPDNTHLTEEVKSSSVSSSSSIDTAQOPSHTTINREBSVQTSDNVEDSHVD 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 115 SVQSTGNKVEVSTAKSDEQASPKSTNEDLNTKQ-----TISQALQ-PDQENKSVYN 168
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 117 FANSKIKESNTSGKEENTIEOPNKVKEDSTSOBSGYTNI-DEKISNDEL-NLPIN 173
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169 ---VOPTNEBKRYD-----AKTEST-----LNVKSDAIKSDELTVDNNSNS 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 174 EYENKARPLSTSAOP-----SIKRVYNQLAEOGSNVHLKIVDOSITEGYDSEGYI 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 210 NNENNADILPKSTAPKRLNTMRIRAAVOPSTTEAKNVNDLTSNTTLTVADAKNKIV 269
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 230 KAHDAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTPIKID-NSG 282
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 270 PAQDYLSLKSOIT--VDKVKSGDYFTIKY-SDTVQVYGLNEDIKN--IGDIDPNNG 323
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 283 ELIATGYDNKKQOITTYFTDVKYENIKAKHLKLSYDKSKVPNNKTLDEYKTAJS 342
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 324 ETIATAKHDTANLLITYFTDYDRFNSVOMGINSIYMDADTIP--VSKNDVEFWTIG 381
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 343 SVNKTITVEYQRPN--ENRTANLOSMTNIDTKNHT-----VEQTIYINPLKRSKE 392
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 382 NNTTATTANIOYPDVYVNEKNSIGSAFT--ETVSHGKNENGYKQTIYVNPSENLTN 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 393 TVNVI-----SGNGDEGSTIIDSTIIKVKVGDNONLPSNRIDYS--EYEDVTNDYA 446
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 AKLKVQAVHSSYPNNIGQINKDVTDIKIYQPKGYTL--NKGVDYNTKELTDVYN-QYL 495
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 447 Q---LGNNDVINFGNIDSPYIIKIVISKYDPNKDDYTTIOQTVMTTINETYGEFRTA 503
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 496 OKITTYGDNNSAVIDFGNADSAVYVNVNTEFOYTNSSEPLVMATLSSTGN-----KSV 549
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 504 SYDNTIAFSTSSGOGDGPPEKTYKIGDYVEDVDKQIONTNDNEKPLSVLTVLTP 563
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 550 STGNALGFTNNOSGAG---OEYKIGNTVWEDTKNKGVOEL--GEKGVNVTYTV-FD 602
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 564 DGTSKSVR---TDEDGY 578
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 603 NNTNTRKVGAVTKEDGSY 620
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

US-09-815-242-13137

; Sequence 13137, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELTRA 011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; APPLICANT: Trawick, John D.

; APPLICANT: Wall, Daniel

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Haselbeck, Robert

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; APPLICANT: Trawick, John D.

```

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13137
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-13137

```

Query Match 19.1%; Score 576; DB 10; Length 1349;

Best Local Similarity 30.1%; Pred. No. 9.8e-24;

Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

```

OY 1 SDEEKNV---INNOSINTDNNQIIEETNNVDIEKRSDETESTVND-NEAT 56
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 55 STNKLNETTASDNOSSDKVDKMOQLNEDNTKNDKEMVSSOGNETTSGNKILIEE 114
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 57 FLQKTPDNTHLTEEVKSSSVSSSSSIDTAQOPSHTTINREBSVQTSDNVEDSHVD 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 115 SVQSTGNKVEVSTAKSDEQASPKSTNEDLNTKQ-----TISQALQ-PDQENKSVYN 168
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 117 FANSKIKESNTSGKEENTIEOPNKVKEDSTSOBSGYTNI-DEKISNDEL-NLPIN 173
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169 ---VOPTNEBKRYD-----AKTEST-----LNVKSDAIKSDELTVDNNSNS 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 174 EYENKARPLSTSAOP-----SIKRVYNQLAEOGSNVHLKIVDOSITEGYDSEGYI 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 210 NNENNADILPKSTAPKRLNTMRIRAAVOPSTTEAKNVNDLTSNTTLTVADAKNKIV 269
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 230 KAHDAENLIYDVFEDVKVSGDTMTVDIDKNTV-----PSDLTDSFTPIKID-NSG 282
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 270 PAQDYLSLKSOIT--VDKVKSGDYFTIKY-SDTVQVYGLNEDIKN--IGDIDPNNG 323
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 283 ELIATGYDNKKQOITTYFTDVKYENIKAKHLKLSYDKSKVPNNKTLDEYKTAJS 342
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 324 ETIATAKHDTANLLITYFTDYDRFNSVOMGINSIYMDADTIP--VSKNDVEFWTIG 381
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 343 SVNKTITVEYQRPN--ENRTANLOSMTNIDTKNHT-----VEQTIYINPLKRSKE 392
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 382 NNTTATTANIOYPDVYVNEKNSIGSAFT--ETVSHGKNENGYKQTIYVNPSENLTN 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 393 TVNVI-----SGNGDEGSTIIDSTIIKVKVGDNONLPSNRIDYS--EYEDVTNDYA 446
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 AKLKVQAVHSSYPNNIGQINKDVTDIKIYQPKGYTL--NKGVDYNTKELTDVYN-QYL 495
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 447 Q---LGNNDVINFGNIDSPYIIKIVISKYDPNKDDYTTIOQTVMTTINETYGEFRTA 503
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 496 OKITTYGDNNSAVIDFGNADSAVYVNVNTEFOYTNSSEPLVMATLSSTGN-----KSV 549
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 504 SYDNTIAFSTSSGOGDGPPEKTYKIGDYVEDVDKQIONTNDNEKPLSVLTVLTP 563
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 550 STGNALGFTNNOSGAG---OEYKIGNTVWEDTKNKGVOEL--GEKGVNVTYTV-FD 602
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 564 DGTSKSVR---TDEDGY 578
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 603 NNTNTRKVGAVTKEDGSY 620
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 3

US-09-815-242-5578

; Sequence 5578, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

```

APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5578
LENGTH: 932
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5578

Query Match      16.6%  Score 499.5; DB 10; Length 932;
Best Local Similarity 27.7%; Pred. No. 8.6e-20;
Matches 160; Conservative 95; Mismatches 212; Indels 111; Gaps 24;

OY 35 GIEKSEDRTESTTVDENEAFLQKTPDNDHLEEEVKESSVSNSIDTAQPSH 94
DB 45 GHEAKAHEHTNGELNOSKNETT-----APSEN-----KTEKVDSRQ--- 82
OY 95 TTINREESVOTSDNVEDSHVSPFANSKIKESNTESGKEENTIEOPKNKVEDSTTSPSGY 154
DB 83 -----NNIEOSTTSD--QPKVNESDNTSVKE--TTEEP---QNTTSTQPTKQ 122
OY 155 TNIDEKISNODLELNPINEYENKARPLSTSAQPSIK-----RYVNOLOAA-EOGSNNY 208
DB 123 NN--DAMAKD--NLAONISTQADKDVSTTPKTTIKRPTLNRMVNVNVAAPQOQTNN 177
OY 209 HLIKVTDOSI-----TEGYDSEGVYKAHDAENLIYDVFVEDKVKSGDTM 255
DB 178 DKVHFTNIDIALDKGHVNKTGTGTEFWATSSDYLK-----LKANYTIDSVKRGDTF 229
OY 256 TVDIDKNVPSDLTDSFTIPKIKDNGSEIATGTYNKKNKQIYTTTVDYDKENIKAKHL 315
DB 230 TFRYGGYFRPGSVRLPSQOTONLYNAOGNIIAKGIYDSESTTYYTFTFNVDQYTNISGSF 289
OY 316 KILSYIDSKSVPNNTKLDVEYKATLSSVKNKTTVEYQRPNEKRTANLQSMFNIDTKNH 375
DB 290 EOYAFKREKRNATTDKATYAPREVTGLNDKYSKNYIVD---GNOKGOQLSSTNYINNEDL 346
OY 376 TVEQTYIYN-PLRYSAKETNV-NISGNGDEGSIIDDSIIKVKYKGDNON-----LPD 427
DB 347 SRMNTYVYVNOPKRTYKKEFTVNLJ-----GYKFNDAKFKIYEV-TNQNGVDSFTPD 400
OY 428 SNRIYDYSEEDVTNDYDQALGNNDVYNIFG--NIDSPYIIVKISKYDENDKDYTTIQ 485
DB 401 TSKLTVDVTKFKIT---YSNDNKTATVDLNGQSSSDKYITQOAVYPDNSSTDNKGIDY 457
OY 486 TYTMQTTINTEYGEFTADNTIARSTSSGQGGOLPEPKTKIKIDYVWEDVDKXION 545
DB 458 TLETQNGKSSWSN-----SYSNVNGSSTANGD-----OKKYNLDYVWEDVDNKGKOD 505
OY 546 TNDNEKPLSNVLTLYTPDQTS-KSVRTDEDGKYOPDG 582
DB 506 A--NEKGIGVYIILKDSNKGELDRITTTDENGKIOFTG 543

```

```

RESULT 4
US-09-815-242-12438
Sequence 12438, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 12438
LENGTH: 932
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12438

Query Match      16.6%  Score 499.5; DB 10; Length 932;
Best Local Similarity 27.7%; Pred. No. 8.6e-20;
Matches 160; Conservative 95; Mismatches 212; Indels 111; Gaps 24;

OY 35 GIEKSEDRTESTTVDENEAFLQKTPDNDHLEEEVKESSVSNSIDTAQPSH 94
DB 45 GHEAKAHEHTNGELNOSKNETT-----APSEN-----KTEKVDSRQ--- 82
OY 95 TTINREESVOTSDNVEDSHVSPFANSKIKESNTESGKEENTIEOPKNKVEDSTTSPSGY 154
DB 83 -----NNIEOSTTSD--QPKVNESDNTSVKE--TTEEP---QNTTSTQPTKQ 122
OY 155 TNIDEKISNODLELNPINEYENKARPLSTSAQPSIK-----RYVNOLOAA-EOGSNNY 208
DB 123 NN--DAMAKD--NLAONISTQADKDVSTTPKTTIKRPTLNRMVNVNVAAPQOQTNN 177
OY 209 HLIKVTDOSI-----TEGYDSEGVYKAHDAENLIYDVFVEDKVKSGDTM 255
DB 178 DKVHFTNIDIALDKGHVNKTGTGTEFWATSSDYLK-----LKANYTIDSVKRGDTF 229
OY 256 TVDIDKNVPSDLTDSFTIPKIKDNGSEIATGTYNKKNKQIYTTTVDYDKENIKAKHL 315
DB 230 TFRYGGYFRPGSVRLPSQOTONLYNAOGNIIAKGIYDSESTTYYTFTFNVDQYTNISGSF 289
OY 316 KILSYIDSKSVPNNTKLDVEYKATLSSVKNKTTVEYQRPNEKRTANLQSMFNIDTKNH 375
DB 290 EOYAFKREKRNATTDKATYAPREVTGLNDKYSKNYIVD---GNOKGOQLSSTNYINNEDL 346
OY 376 TVEQTYIYN-PLRYSAKETNV-NISGNGDEGSIIDDSIIKVKYKGDNON-----LPD 427
DB 347 SRMNTYVYVNOPKRTYKKEFTVNLJ-----GYKFNDAKFKIYEV-TNQNGVDSFTPD 400

```



```

OY 91 OPHSTTTINREBSVQTSNDVEDSHSDVANSKIKESNTESEKKEENTIBQPKVKEDSTSO 150
Db 93 QPKVT-----MSDAATYKETSNNOS-----PONATANOSTIK 125
OY 151 PEGYTNIDEX---ISNODELLNPEINERYENKARPLSTTSAOSPIK-----RYTVNOLOA- 201
Db 126 TSNVTTNDKSTSTYSNTEIDKSNL-----TQAKDVSTPRTTIKPRTLNRMAVNVTAAP 179
OY 202 EOGSNVHLIKVT--DOSITEGY--DBSEGYK--AHDAEMLIDYTFEVDKYSQDTMT 256
Db 180 QPGTNVNDKVFHNSNIDAIKIDKGHVNOQTGKTEFWAASSDVLKUKANYTTIDSVKEGDTFT 239
OY 257 VDDIKNTVPSPDLMDSFTTPIKIDKSGEITLNGVTDNKNKOITFTFTDYKYEIKAHLK 316
Db 240 FKIGQYRPGSVRLPSTOTONLYNAGNINIKKCIYDSTNTTITFTYKVOYTRNRSFE 299
OY 317 LTSYIDSKVPNNNTKLDVEKTKALSSSVNKTITVEYQRPENRNTANLQSMFTNIDRKNT 376
Db 300 QVAPAKKRNKNTTOKTAVKMEVTLGNDPYSEIYDV---GNKKAQPLISSTNYINNEDLS 356
OY 377 VEOITTYIN-PLRYSAKETNV-NISGNDEGSTIIDSTTIKVKYGVGNOL-----PDSN 429
Db 357 RNMAYVNOPKNTYTKOTFEVNTLP-----GKFNPNKRNKRIYEVTDONQVDSFEPDTS 411
OY 430 RIVYSEVEDTNDYDQOLGNNNNVNFNGNIDS--PYIIKVI-----SKYDPNKDDYTT 482
Db 412 KUKVTOFOFDV---YSNDKNTATVYDLKMGOTSSNKQITIQOJYAVPNNSSTDGKIDY-- 466
OY 483 IQGTVTMOGTINETYGEFRFASDYNTIAFTSSGOGGDDLPPEKTYKIGDYVEDVDKDG 542
Db 467 -----TLDTOKTKYSW---SNSYVNVGSSSTANGD-----QKYNMIGDVWEDTKKG 511
OY 543 IQNTNDEKPLSNVLVLTITYDGS-KSVRPDEGKQFOFG 562
Db 512 KODA--NEKGIGYVYILKDSNGKELDRTTTIDENGKYQFOFG 550

RESULT 7
US-09-815-242-5471
Sequence 5471, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5471

```

```

      LENGTH: 1021
      TYPE: PROT
      ORGANISM: Staphylococcus aureus
US-09-815-242-5471

Query Match          14.9%; Score 449; DB 10; Length 1021;
Best Local Similarity 24.7%; Pred. No. 5,2e-17;
Matches 159; Conservative 115; Mismatches 247; Indels 124; Gaps 28;

OY       2 SDEEKNDVI-----NNNQSIINTDNNQI IKKEFTNMYDGIEKRESDERTESTYNVDENEA 55
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       37 ADASENSVTQGDASANESEKSSDDSSVNAAPKTDNTNVSD-----TKTSTNNNGET 87

OY       56 TFLQKTPODDNHLTEEEVKESVSNNSSIDTAQP----SHTINNEESQTSDNVD 111
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       88 SVAQNPAAQET-----TQASTNAVTEETPYTGAEATYTATQAANPATQOSSN 135

OY       112 SHVSDFANSKIKESNTESGKEENTIEQPNKVXEDSTTSOPSGTYINDEXISMODELL-L 170
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       136 TNAELVN---QTSNETASNDINIY-----SSVNSPONTAKENSTQDSTEXT 183

OY       171 PINEENKARPLSTSAOPSIRKYTN-----QLAA-----EQGSNNHLLIKYVD 215
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       184 PSN---NSAQSOSTASKNKDYNQAVNTPAPRMAFSLAAVAADAAPAQDTITN--QLTD 238

OY       216 QSITEGYDDEGVKAHADENLIDYTEVEVDKVASGDTMYVIDKNIYPSDLTDSFTIP 275
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       239 VKVT---IDSGTTVVPHQAGYVKLNYGFSPNSAVKCDTFKITVPFELELNIGVSTAKVP 295

OY       276 KIKDMSGELIATGTVDNNKKOTITFTDYVDKYEENIKAHILKLTYSYIDSKVPPNNNTKL-D 334
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       296 PLMGD-OVLANGVIDS--DNVITYFTFDVDRKENVTANTTPAIID----PENVTKTGN 349

OY       335 VEKPTAL--SVSNKTIITEYQRPNERNRTANLOSMFTNIDTKHHTEVOTIYNPLRYSAKE 392
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       350 VLTTCIGCTNMTASKVLIDYEKGQFNHSIKGLTDIQDKTNTNTYQTLIVNP-----SG 404

OY       393 TNV-----NISGNDECSFTIIDSTIIKYKVGDNQMLPDSNRJIYDSEVEDYTNDDY 445
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       405 DNVTVPALTLGNLIPTMKSNALLIDAKNTDIKYRV--DNANDLSesyvvnpsDEDYTNQVR 463

OY       446 AOLGNNDVINMF---GNIDSPYIIKIYISKDPKNKDOTTQQFTVMQTLTNEYGEF- 500
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       464 ISEPANAQYKVEFPDDDOITTPYLVVGHIDPASTG-----DIALKSTFYGDSNF 517

OY       501 -RTASVDNTIAESTSGSQGG-Q-DLP--PEKTYKIGDY--VMEVDYKD----- 541
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       518 WMSMWDNVAANNGSGSDGIDKPVYPQPDPEGIEPIEDSDSDDPDSDSDSDSDS 577

OY       542 ---GIQNTNDEKPLSNVLVTLTYYPDGTSKSVRT--DEDCQIOFD 581
           :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB       578 GSDSGSDSTSDDSASDSDSDSDSDSDSDSDSTSDSDSASDSD 622

RESULT 8
US-09-815-242-12544
Sequence 12544, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA_011A
CURRENT APPLICATION NUMBER: US/09/815_242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078

```


QY 485 QVTMOTINETYGEFRTA--SYDNITAFSTSSGOGG-----DLPEKTYKIG 531
 DB 440 --LEFQTHLEGYNNYTTNLTWKNGVAFYSNNAOGDKLKEPIIEHSTPIELEFKSE 497
 QY 532 DYVMEVDVDKGIQNTNDNEKPLSNVLTLT-----YPDGTSKSVRTDED 575
 DB 498 PVEKHELTGTIEESNDS-KPIDEFYHTAVEGAEHAGET---IETDED 542

RESULT 10

US-09-815-1242-13140
 Sequence #3140, Application US/09815242
 Patent No. US20020061569A1

GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 13140
 LENGTH: 767
 TYPE: PRT
 ORGANISM: Staphylococcus aureus
 US-09-815-242-13140

Query Match 12.5%; Score 376; DB 10; Length 767;
 Best Local Similarity 23.4%; Pred. No. 3.2e-13;
 Matches 18; Conservative 110; Mismatches 217; Indels 124; Gaps 22;

QY 35 GIKRSRDRSTETTNVDENATFLQKPPQNTHTLEEVESSSVSSNSIDTAQ---- 90
 DB 30 GQEKAAASDQNTTVEES-----GSSATESKASHTQTTNNVNTIDETQSYSA 78
 QY 91 ----OPSHHT-INREESVOT--SDNVEDSHVDPFANSKIKESNTESGKEETIOPKVK 143
 DB 79 TSTEDPSQSTOVVTEEPKPVQAPKVTSTRY-DLPSEKVALKETTGTQVD--IAQPSNVS 135
 QY 144 EDSTTSPSGYTNIDKISNODLLNLPINEYENKARPLSTSAOPSIKRYVNOQLAAQ 203
 DB 136 E-----IKPRKRSITDVTAAEK 153
 QY 204 -----GSNNHLIKVTDOSTTEGYDDESGYIKAHDAENLIYDVTFEYDDKVKSGDT 254
 DB 154 EYVEETKATGTDTNKKVEEGSEIYGHKQDTNVNPNHMERVTLKRYKMGEGIKAGDY 213
 QY 255 MYVDIDKNTVPDLTDSFTTPIKIDNGSEIIATGTYNKKQOLYTFYTDVYDKENIKAH 314
 DB 214 FDTTLSQNVETHGISLIRKVPKIKSTIDGQVMAIGEITIGERK-VAYITKEVYOEKKDLTAE 272

QY 315 LKLTSTYIDSKVPPNNNTKLDVEKKTALSSVNTKITVEY---ORPNENRANLQSMFTNID 371
 DB 273 LSLNLFDPTTYQKQNO--NVEYKLGFTYVSKFENIQYLAGVDNNGVTAN-----GRID 326
 QY 372 TKNHT---VEOTIYINPLRYSAKETNVNISGNDDEGTTIIDSTIIKRYKVGDNOLPDS 428
 DB 327 TLNKVQGKSFHFAYMKPNNQSL--SSVTVYQGTGKNGKPPVNNPTVKYVYKHISDDLAE 384
 QY 429 --NRIYDSEYEDVTDVYAQLGNNDVNINFGNIDSP--YIIRKVISKIDPNKDDYTTIQ 484
 DB 385 VYAKLDVSKFEDVTFNMSLDFPTNGCYSLNPNLDQSKNYVIKYGYSNASN----- 439
 QY 485 QVTMOTINETYGEFRTA--SYDNITAFSTSSGOGG-----DLPEKTYKIG 531
 DB 440 --LEFQTHLEGYNNYTTNLTWKNGVAFYSNNAOGDKLKEPIIEHSTPIELEFKSE 497
 QY 532 DYVMEVDVDKGIQNTNDNEKPLSNVLTLT-----YPDGTSKSVRTDED 575
 DB 498 PVEKHELTGTIEESNDS-KPIDEFYHTAVEGAEHAGET---IETDED 542

RESULT 11

US-09-813-820-7
 Sequence 7, Application US/09813820
 Patent No. US20020102262A1

GENERAL INFORMATION:
 APPLICANT: Hook, Magnus
 Patti, Joseph M.
 House-Pompeo, Karen
 Stahanam, Narayana
 Symersky, Jindrich
 TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: U.S.
 ZIP: 77210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/813,820
 FILING DATE: 22-Mar-2001

CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/856,253
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Kitchell, Barbara S.
 REGISTRATION NUMBER: 33,928
 REFERENCE/DOCKET NUMBER: TAMK.193
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 345 amino acids
 TYPE: amino acid
 STRANDEDNESS: <unknown>
 TOPOLOGY: linear

US-09-813-820-7
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 11.6%; Score 349; DB 10; Length 345;
 Best Local Similarity 30.8%; Pred. No. 3.6e-12;
 Matches 102; Conservative 61; Mismatches 136; Indels 32; Gaps 15;

[illegible]

[illegible]

```

1 PRIOR APPLICATION NUMBER: 60/206,848
2 PRIOR FILING DATE: 2000-05-23
3 PRIOR APPLICATION NUMBER: 60/207,727
4 PRIOR FILING DATE: 2000-05-26
5 PRIOR APPLICATION NUMBER: 60/242,578
6 PRIOR FILING DATE: 2000-10-23
7 PRIOR APPLICATION NUMBER: 60/253,625
8 PRIOR FILING DATE: 2000-11-27
9 PRIOR APPLICATION NUMBER: 60/257,931
10 PRIOR FILING DATE: 2000-12-22
11 PRIOR APPLICATION NUMBER: 60/269,308
12 PRIOR FILING DATE: 2001-02-16
13 NUMBER OF SEQ ID NOS: 14110
14 SOFTWARE: FASTSEQ for Windows, Version 4.0
15 SEQ ID NO 12838
16 LENGTH: 1018
17 TYPE: PRT
18 ORGANISM: Staphylococcus aureus
19 -09-815-262-12838

```

Query Match	10.2%	Score 308.5	DB 10	Length 1018
Best Local Similarity	21.3%	Pred. No. 2e-09		
Matches 124	Conservative 118	Mismatches 242	Indels 97	Gaps 25

```

0Y 35 GIERKSEBERTSTNNVENEATFLQKTPODI-----TULTEEVKESSVSSSSSIDTAQ 90
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 30 GODEMAAASEKTTTVEEN-----GNSATDNKTSJETOTTATNVHIEFOSYNATV--TE 82
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 91 QPSHTT-INRBE---SVQSDNVEDSHVSDPFANSKIESNTSGREBENTEOPRKVEDS 146
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 83 QPSNATQVTTTEAPRKAVQAPOTQAPANIEFVKEEVA-----EAKQVKEFT 130
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 147 TTSPQSG---TYNIDEKISNDELLNLPINIEKAPKAPLSTSAQPSIKRYT-----V 136
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 131 QSDNSGQORQVDLTPKATONOVAETQV---EVAQPRTASEKPRVTRSADVAEAKEA 186
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 197 NQLAAEQSNVNHILIKYTDOSITEGYDSEGVKAHAENLIYVTFEVDVKYSGSGTMT 256
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 SNAKVEITDVTSKAYVEIGSI -EGHNNTKV-EPHAGORAVLYKKYKAFENGHOGDYFD 244
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 257 VDIKNTVPSDLTDSFTPIKIDNS---GELIATGYDNKNKQIYTFETDYVDYKENT 311
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 245 FTLSNVNTTGVGSNARKRPEIKNSVMAATGEVEGG-----KIRYTFNDEKQVDV 297
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 312 KAHKLITSYIDSKVPPNNNTKLVDEYKTAUSSVN---KTIYEVORPENRNTANLOSM 366
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 298 TAELEINLFIIDPKYVOTNGN-----QITISTLEBQTSKELDYKYKDQIGNYAANLNGS 351
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 367 FTNIDTKHTEVEQTIYINPLKYSAKETPNVNSGMDGSGTITDSTIIKYYK -GQDNOL 423
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 352 IETFEKANNRSHVAFLIP--NNGTTSVTVTGLTMKGSNONGQOPKVRIFEYJNEDI 409
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 426 PDSNRIRY---DYSEYEDVTDNDYDAOLG--NNNDVNINFGNIDEPYIIKYSKYDPNKD- 478
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 410 AKS--VYANTJTDTSKFEKVTSMNSGNLNLQNGSYSLNIEHLDKTYVYHHDGELNLTDE 467
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 479 -DYTIQOTVTMOTTINETYGEFPRTASTDNTIASFSTSGGQGFLLP-----EKTYIGD 533
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 468 VDFRT-QWVGHPREQLYKYYDGRGYTLTMDNGLVYUSKNAMNGKNGCPIONNKEFYK-- 523
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
0Y 533 YVMEVDKDGIONTNDNKEPLSNVLTYLTYPDQTSKSRPD 573
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 534 ---EDTIKETILGQYDKN-----LVTTVEEYDSTLDD 555
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

Search completed: March 14, 2003, 13:07:57
Job time : 12.6918 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01 ; Search time 11.8208 Seconds

(without alignments)
4733.207 Million cell updates/sec

Title: US-09-147-405b-13

Perfect score: 3013

Sequence: 1 SSDEKNDVYINNNOISINTDNOIIRKKEETNNVDGIEKSEDETESTTVNDENAEATFLOK

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR-73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 T30214	fibrinogen-binding
2	1144.5	38.0	1166	2 T28680	hypothetical protein
3	1137.5	37.8	1141	2 E89824	hypothetical protein
4	576	18.1	1315	2 T28679	hypothetical protein
5	559	18.6	1385	2 D89824	hypothetical protein
6	489	16.2	953	2 C89824	hypothetical protein
7	452	15.0	933	2 S41539	hypothetical protein
8	452	15.0	989	2 D89852	fibrinogen-binding
9	397.5	13.2	877	2 F90070	Clumping factor B
10	376	12.5	940	2 S19702	fibrinogen-binding
11	344.5	11.4	961	2 G90053	hypothetical protein
12	326	10.8	1038	2 H90053	hypothetical protein
13	300.5	10.0	1018	2 A32192	fibrinogen-binding
14	239	7.9	1039	2 T30856	protein F2 - Strept
15	225	7.5	1463	2 T30290	AAS surface proteol
16	216.5	7.2	1859	2 AG1085	hypothetical protein
17	216.5	7.2	3394	2 T18501	hypothetical protein
18	212.5	7.1	1125	2 E90598	membrane nucleas
19	211.5	7.0	2206	2 G71611	hypothetical protein
20	208.5	6.9	1072	2 A86827	hypothetical protein
21	208	6.9	1127	2 T28317	ORF MSV156 hypote
22	208	6.9	3724	2 T18427	hypothetical protein
23	207.5	6.9	1711	2 T18429	hypothetical protein
24	207.5	6.9	2401	2 T28676	thoxy protein -
25	207	6.9	769	2 F89870	serine proteinase
26	207	6.9	1939	2 T18372	repeat organellar
27	207	6.9	4550	2 T18440	hypothetical protein
28	206.5	6.9	821	2 S67087	hypothetical prote
29	206	6.8	1302	1 J66009	surface-located me

30	205.5	6.8	1365	2 T30822	
31	205.5	6.8	2510	2 T28160	
32	205	6.8	2481	2 D90011	
33	204	6.8	3844	2 T18402	
34	203	6.7	537	2 A23770	
35	202.5	6.7	1997	2 F71607	
36	202.5	6.7	2269	2 T28677	
37	202	6.7	1093	2 T18275	
38	201.5	6.7	4688	2 F82885	
39	199	6.6	665	2 B71609	
40	198.5	6.6	807	2 B71605	
41	198.5	6.6	1805	2 T02712	
42	198.5	6.6	3147	2 T18674	
43	198	6.6	1308	2 E71622	
44	197	6.5	940	2 A01374	
45	197	6.5	2523	2 T18477	

ALIGNMENTS

RESULT 1					
T30214					
fibrinogen-binding protein - Staphylococcus epidermidis					
C:Species: Staphylococcus epidermidis					
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000					
C:Accession: T30214					
R.Nilsson, M.; Frykberg, L.; Flock, J.I.; Pel, L.; Lindberg, M.; Guse, B.					
Infect. Immun. 66, 2666-2673, 1998					
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.					
A:Reference number: J20781; MUID:98261511; PMID:9596732					
A:Accession: T30214					
A:Status: preliminary; translated from GB/EMBL/DBJ					
A:Molecule type: DNA					
A:Residues: 1-1092 <NLI>					
A:Cross-References: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAN76638.1					
Query Match					
Best Local Similarity 100.0%; Score 3013; DB 2; Length 1092;					
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	SSDEKNDVYINNNOISINTDNOIIRKKEETNNVDGIEKSEDETESTTVNDENAEATFLOK	60		
DB	75	SSDEKNDVYINNNOISINTDNOIIRKKEETNNVDGIEKSEDETESTTVNDENAEATFLOK	134		
QY	61	TPDQNTHLTEEEYKSSVESSNSIDTQOOPSHTTINREESYQTSNDVEDSHVDFANS	120		
DB	135	TPDQNTHLTEEEYKSSVESSNSIDTQOOPSHTTINREESYQTSNDVEDSHVDFANS	194		
QY	121	KIKESNTESGKEENTIEQPKYKEDSTQSPGYTNIDKISNODLNLPINEYENKAR	180		
DB	195	KIKESNTESGKEENTIEQPKYKEDSTQSPGYTNIDKISNODLNLPINEYENKAR	254		
QY	181	PLSTSAQPSIKRYTVNOQLAEGSVNHLIKTQDSITREGYDSSGVIKAHAENLIYD	240		
DB	255	PLSTSAQPSIKRYTVNOQLAEGSVNHLIKTQDSITREGYDSSGVIKAHAENLIYD	314		
QY	241	VFEVDDKYSAGDTMTVDIDKNTVPSDLTGFSPRIKQDSGLITGTVDNKNKOITYT	300		
DB	315	VFEVDDKYSAGDTMTVDIDKNTVPSDLTGFSPRIKQDSGLITGTVDNKNKOITYT	374		
QY	301	FTDYVDKYNIAKHLKLTYSIDKSKVPNNNTKLDVEYKATLSSVNTTIVEYQRPENRT	360		
DB	375	FTDYVDKYNIAKHLKLTYSIDKSKVPNNNTKLDVEYKATLSSVNTTIVEYQRPENRT	434		
QY	361	ANTQSMFTNIDTNHNYEQITINPLRYSAKKEFNVNSGNGDGSFTIDSTIIKYYKVG	420		
DB	435	ANTQSMFTNIDTNHNYEQITINPLRYSAKKEFNVNSGNGDGSFTIDSTIIKYYKVG	494		
QY	421	DNQNLDPDSNRITYSEVEDTNDQYQOLGNNNDVNIFFGNDSPYIIKVIKSYDPKNDV	480		
DB	495	DNQNLDPDSNRITYSEVEDTNDQYQOLGNNNDVNIFFGNDSPYIIKVIKSYDPKNDV	554		

[illegible]

Query Match 15.0% Score 452; DB 2; Length 933;
Best Local Similarity 26.6%; Pred. No. 2,1e-12;
Matches 154; Conservative 103; Mismatches 224; Indels 86; Gaps 26;

OY 1 SSDEKNDVNNOSINTDNNQIIKKETNNYDGIEKRSEDRTSTTNDENAEFLQK 60
Db 1 SNEKSNDSSVSAPAKRDTDTNV-----SDTKTSNTN---NGETSVAO 91
OY 61 TPQDNTHLTEEVEKSESVESNSSIDTAQPSPHTITRESVOTSDIVESSHVDPAFNS 120
Db 92 NP-----AQDETQSSSTNAATEETEPVTGATTTTNQANTPATQS-SWTNNEELVN- 143
OY 121 KIKESNTGESKEENTIEDPNRKEDSTPSOPSGYTINDEKISNODELIN-LPINEYENKA 179
Db 144 -QINSNETTFMDTNY-----SSVNSQNOSTMAENVSTODTTEATPSN---NES 189
OY 180 RPLSTTSAOPSIRKYTVN-----QLAEOGSNNHLKVTDQ--SITEGYDSEGV 228
Db 190 APOSTDAENKDYNQAVNTSAPRRARAFSLAAVADAPAAGTDITNQLTNVTVGJ-DGGTY 248
OY 229 IKADHAELIYDYFEVDDKRYSGDTMIVDDDKMTVPSDLTDSFTPIKINSEIITATG 288
Db 249 VYPHQAQGVKNLYNGEESPNSAVKGDTEKITVPEKELNLGVSTKVPIMAGD-QVLANG 307
OY 289 TYDNKNKOITTFETDYDVKYENIKAHLEKLTYIDSKVPNNNTKL-DVEYKTALSS--VN 345
Db 308 VIDF-DGNVIYTFPDYVTKDVKATILMPAYID----PENVKTKGNTLATIGGSTIAN 362
OY 346 KTIVEXYORPENRNTANLOSMFTNIDTKNHTVEQTITINPL-RRSAKETVNIISGNDE 403
Db 363 KTVLEVDEYKGFYLLSIKGTIDQIDKTNNTYROTIVYPGSDNVIAFVLGNLKPNTDS 422
OY 404 GSTIIDSTTIKKYKVGONULPNSNRIFYDSEVEDYTNDDYAOLGNNNDVINFG---- 459
Db 423 NALIDQONTSIKRVKVDNAADLESYFY-NPEDEFEDYNSVNIIFPNNOYKVEENFPDD 481
OY 460 NIDSPYIKIVSKSYDPN-KDDYTTIQQTVMQTITNEYTGE--FRITASDYNTIAFSTSSG 516
Db 482 QITPEYIYVNGHIDPNKSGD-----LAIRSTLYGNSNIIMRSMWMDNEVAFNNGSG 534
OY 517 OGQG-DLP--PEKTYIKIGDY--WVEDVDKD-GIONTFDN 549
Db 535 SGDGIKDFVPEQDPDEPCEIPEDSDSDPSGSGSDS 573

RESULT 8
DB9852
fibrogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (str
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence-revision 10-May-2001 #text-change 22-Oct-2001
C:Accession: D89852
R:Kuroudo, M.; Ohts, T.; Uchiyama, I.; Baba, T.; Yusaawa, H.; Kobayashi, I.; Cut, L.; O
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant staphylococcus aureus.
A:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: D89852
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-989 <KJR>
A:Cross-references: GB:BA000018; PID:g13700678; PION:BAB41975.1; GSPPDB:GN00149
A:Experimental source: strain NJ15
C:Genetics:
A:Gene: cflA

Query Match 15.0% Score 452; DB 2; Length 989;
Best Local Similarity 25.9%; Pred. No. 2,3e-12;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

OY 1 SSDEKNDVNNOSINTDNNQIIKKETNNYDGIEKRSEDRTSTTNDENAEFLQK 60
Db 1 SNEKSNDSSVSAPAKRDTDTNV-----SDTKTSNTN---NGETSVAO 91

```

Db 51 SNEKSNDSSSVSAAPKPTDDTNV-----SPTKTSNTN---NGETSVAQ 91
Qy 61 TPQDNTHLTEEEVKESSSVSSNSIDTAQOPSHTTINEESVQTSNDVEDSVSPFAS 120
Db 92 NP-----AQOETTSSTNATTEETPVNGEATTTTNOAMPATQOS--SNTNAELVN- 143
Qy 121 KKESTESGKEENTIEQNKVKEDSTTSQPSGYTIDEXISODELLN-LPINEYENKA 179
Db 144 --QTSNETTSNDNTV-----SSVNSPONTNAENVSTQDTSTATPSN---NES 189
Qy 180 RELSTTSA-----OPSIKRVYNOLA-----EOGSVNHILIKYTDOSTEGYD 224
Db 190 APOUNTDAKQDVYSAOVNSTPRMRFLSLAAVAADAPAGTDTN--QLTDVKVT---ID 244
Qy 225 SEGVIKHADEMLIYDVTEFVDKYSQDTMTVDIDKNTVPSDLTDSFTPIKIDSGEI 284
Db 245 SGTIVYPHOGYVKLNGVGSVPNSAVKGDPTKIVPEKELNLTNGVSTAKVPPIMAGD-QY 303
Qy 285 IATGTYDNKKNQITTYFTFYVDKYEKIKAKLSTYIDSKVYNNNTKL-DVEYKTAAL-- 341
Db 304 LANGVIDS--DGNVYTFYFDYVDNKENVTANTIMPAYID---PENVTKGWVLTGTIGT 358
Qy 342 SSVNKTIVYQORPNENRANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNV----- 395
Db 359 NTAASKTVLIDYERKGFHLSIKGTIDQIDKNTTYRQITVYVNP-----SGDKVVLPAUL 413
Qy 396 -NISGNGDEGTYIIDSTIIKYYKVGDNQNLPSNRIDYSEXEDYTDNDYAOLGNNNDY 454
Db 414 GNLIPTKSNALIDAKNTDKIVYRV--DNANDLESYYVNSDPEDEYTNQVIRSFPAANOY 472
Qy 455 NINF-----GNISPYIIKISKYDPKDKDYTTIQOTVMTQITNEYTGER--STASYDNT 508
Db 473 KVEEPTDDQITTPYIVVNGHIDPASTG-----DLAKSTYIGDSNITWMSMDNE 526
Qy 509 IAFSTSGOGQOG-DLP--PEKTYKIDY--VMEVDVKD-GIQMTNDN 549
Db 527 VAFNNSGSGDGDIDKVPVEQPPPEIEPIEPDSDSPOSDSGSDS 573

RESULT 9
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hatori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; M0ID:21311952; PMID:11418146
A:Accession: F90070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <R>
A:Cross-references: GB:BA000018; PID:g13702588; PIDN:BA843728.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 13.28; Score 397.5; DB 2; Length 877;
Best Local Similarity 24.5%; Pred. No. 4.3e-10;
Matches 143; Conservative 81; Mismatches 207; Indels 153; Gaps 23;

```

```

Db 143 NSQVDNKT-----NDANSIATNSELKNSQT-----LDLP-----Q 173
Qy 179 APRLLSTSAQ-----PSIKRVYNOLAEO-----GSNVNHLIKYTDOSTEGYD 225
Db 174 SSPORTISNAQGSKSVRRAVRAVSLAVAPVYNAADAKGTNN-----DKYASNFKE 227
Qy 226 EGYKHADEMLIYDVTEFVDKYSQDTMTV-----DID-----KNVPSDLTDS 271
Db 228 KTFEPDNGSGTFMAANFTVTDKVSQDYFTAKLPDLSLNGNDVDYSNNTNP----- 281
Qy 272 FTPIKIDSGEIIATGYDNKNKQITTYFTFYVDKYEKIKAKLSTYIDSKVYNNNT 331
Db 282 --IAQKSNNGDVYAKATDILKTYTFVFTOVVNNKENINQGSFSLPFTDRAKAKSG 339
Qy 332 KIDVEYKTAALSSVNKTIVYQOR-----NENRANLOSMTNIDTN--HYEQTIYIN 385
Db 340 -YDANINIDEMFNKITYNYSPIAGIDKPNGANISOIIGVDTASGONTYKQTFVNP 398
Qy 386 LRSAKETVNVISGNDGTYIDS-----TIKYYKVGDNQNLPS----- 428
Db 399 KORVLGNTWYIYIKGYDK-----LESSGKVSATDTKLRIEVDNYSKLSDSYIADPNDS 454
Qy 429 -----NRIYSEYEDVTNDYAOLGNNNDVINFGNIDSPYIIKIVSKYDPNKD 479
Db 455 LKEVTQFKNRIY-----YE-----HPVVASIKGDDIKTYVVLVEGHYDNTGKN 499
Qy 480 YTT--IQOTVMTQITNEYTGERFASDYNTIAFSTSGOGQD 521
Db 500 LKTOVIOENVDVPT-----NRQYSIFGMNENNVRYVGGSGADSD 538

RESULT 10
S19702
Fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S19702
R:Joensson, K.; Stignaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus
A:Reference number: S19702; M0ID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <O>
A:Cross-references: EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g581562
C:Keywords: fibronectin binding

Query Match 12.58; Score 376; DB 2; Length 940;
Best Local Similarity 23.48%; Pred. No. 4e-09;
Matches 138; Conservative 110; Mismatches 217; Indels 124; Gaps 22;

```

[illegible]

Db 506 --KNGQITQDNDPEKEDTAKGTMSGQYDAKQITETENODNTPIDIDYHRAIDEGGX- 562
 QY 580 FDG 582
 Db 563 VDG 565

RESULT 13

A32192
 fibronectin-binding protein - Staphylococcus aureus
 C:Species: Staphylococcus aureus
 C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
 C:Accession: A32192
 R:Signaes, C.; Raucel, G.; Joensson, K.; Lindgren, P.E.; Anantharamiah, G.M.; Hoeck, M.
 Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989
 A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylo
 A:Reference number: A32192; MUID:89098998; PMID:2521391
 A:Accession: A32192
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1018 <SIG>
 A:Cross-references: GB:J04151
 C:Keywords: fibronectin binding

Query Match 10.0%; Score 300.5; DB 2; Length 1018;
 Best Local Similarity 21.2%; Pred. No. 8.1e-06;
 Matches 123; Conservative 118; Mismatches 245; Indels 97; Gaps 25;

QY 35 GIEKSEDETESTVDNEATFLOKTPDN-----THLFEVYKSSSVSSSSSIDTQ 90
 Db 30 GQDKRAAASEQKTTVEEN-----GNSATDNKTSSETQTTATVNHIEQTQSYNAV--TE 82
 QY 91 QPSHTT-INREE---SVQTSNDVEDSHVDFANSKIKESNTESGKEENTIEQPNKVEDS 146
 Db 83 QPSNATQVTEAPKAVQAPQTAQANITETVKEEYKE-----EAKPYKERT 130
 QY 147 TTSQPSG---VTNIDEKISNDELNLPLINEXENKARPLSTSAQSIKRV-----V 196
 Db 131 GQDONSQDQRYDLPRKKAQNOVAETQY---EVAQPTASESKPRVRSADVAEAKBA 186
 QY 197 NOLAEQGSNNVHLKVTQOSTEGYDDESGYKADANLYDYTFEVDKVKSGDMT 256
 Db 187 SNAKVEGTGVTSKYAVEIGSI-EGHNNTNKV-EPHAGRAVLKYLKFEENGJHQGDYD 244
 QY 257 VOIDKTVPSDLTSEFTIKIKDNS---GELIATGYDNKKQITTYFTDVVKYENI 311
 Db 245 FTLSNNVNHGVTAKKVEIKNGSVYMATGEVLEGG-----KIRYFTNDIEDKDV 297
 QY 312 KAHLKLTSTYIDSKVYNNNTKLDVEYKTLASSVN---KITVEYQRPNEKFTANLQSM 366
 Db 298 TALEELNLFIDPKVTQTN-----QITSTLINEQTSKELDKYKDGIGNYANLNGS 351
 QY 367 FENIDKKNITVEQTIYNLAKRSKFTNNISGNGDEGSIIDDSITIKYK- VGNOML 425
 Db 352 IETFNANRRFHVAFIKP--NNGKTSVTVTGTLLKNGSNQNGQPKRVKIFELGNENDI 409
 QY 426 PSNRIRY----DYSEYEDTNDYAQLG--NNNDVININGNIDSPYIIKIVSYDNKPD 478
 Db 410 AAS--VYATMTDTSKKEVTSNMSGNLNIONNGSYSLNINENIDKTYVHVHDSGYLGTBE 467
 QY 479 -DYTTIQQTVMQTTINETYGEFRTASYNQNTIAFSTSSQCGDLP-----EKYKIGD 532
 Db 468 VDFRT-QMVGHPQLKYYDYDRGYTLTMNGLIVLNKNGKNGKNPIIONNFEYK--- 523
 QY 533 YVMEVDKXGCIQNTNDNEKPLSNVLYTLTYPDGTSKSVTD 573
 Db 524 ---EDTKEFTLGGYDKN-----LVTVVEEYDSSTLDD 555

RESULT 14
 T30856
 protein P2 - Streptococcus pyogenes
 C:Species: Streptococcus pyogenes

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T30856
 R:Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanskl, E.
 Mol. Microbiol. 21, 373-384, 1996
 A:Title: Protein P2, a novel fibronectin-binding protein from Streptococcus pyogenes
 A:Reference number: 220907; MUID:97011581; PMID:8858591
 A:Accession: T30856
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1039 <JAF>
 A:Cross-references: EMBL:U31980; NID:g1654115; PID:g1654116; PIDN:AAC44522.1
 C:Genetics:
 A:Note: PRPF2

Query Match 7.9%; Score 239; DB 2; Length 1039;
 Best Local Similarity 22.6%; Pred. No. 0.0038;
 Matches 93; Conservative 79; Mismatches 160; Indels 72; Gaps 18;

QY 217 SITGEGDSEGVYKADH---AENLIYDVFEDVKVKS-----GDTMTVDIDKNTVPSDL 268
 Db 38 SVSKYGEQEKTSNADYRNHAAVEKMSFELKQKDKSETINPGDTEFLQIDRLNPKGI 97
 QY 269 TDSFTPK-IKDNSEIATGYDNKKNQITTYFTDYDKYENIKAKLSTYIDSKVY 327
 Db 98 SOD--TPKITDSENSPLAIGKTDKTHQITTYFTNYTAGLKVLSAELSLEKLEYL 155
 QY 328 NNNTKLDVEYKATLSSVKNITVEYQRPNE-----NRTANLQSMFTNIDTKHVE 378
 Db 156 ENTNIDSDFKSTIGGEITIKYKGVNVLYGNSKESNYITNGLSNVGGSIESVNTETGEV 215
 QY 379 QTIYNPLRYSKETNNVNSG-----NGDESTIIDS--TIVYKGNQNPDSNR 430
 Db 216 WYVYVNPNTNIPYAVNLMLGFAKRTAOGENNSLYVSQOLGYDIEYPPHYRLPTSYG 275
 QY 431 IYDSEYEDVTNDVYAOI-----GNNDVNINFG-NID-SPYIIKIVISKYDPNKDYTT 482
 Db 276 V-DISRL-NLRKDLEKLPQSGTQGANRRLIDFGENLQKAFVYKVGKADQSGKE--- 330
 QY 483 IQQVTMTQITNEYT--GEFRT-----ASYDNTIAFSTSSQCGQ-----DLPEPT 527
 Db 331 ----LIVQSHLSFNNMGVYKTLRPNSHVSFNEIALSPSKSGSSTSEFTKPTIVANL 386
 QY 528 YKIGDYVEDVDKDG-----QNTDNEKPL---SNVLYTLTYPGTS 567
 Db 387 KRVQALRFKKVSTDNVPLPEAFELRSSNGNSOKLEASSNTQGEIHFQDLTS 438

RESULT 15

T30290
 AAS surface protein - Staphylococcus saprophyticus
 C:Species: Staphylococcus saprophyticus
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T30290
 R:Hell, W.; Meyer, H.G.W.; Gattermann, S.G.
 Mol. Microbiol. 29, 871-881, 1998
 A:Title: Cloning of aas, a gene encoding a Staphylococcus saprophyticus surface pro
 A:Reference number: 220809; MUID:98389318; PMID:9723925
 A:Accession: T30290
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-1463 <HEL>
 A:Cross-references: EMBL:AJ000007; NID:e1295679; PID:e1295630; PIDN:CA03852.1
 C:Genetics:
 A:Gene: aas

Query Match 7.5%; Score 225; DB 2; Length 1463;
 Best Local Similarity 20.9%; Pred. No. 0.024;
 Matches 143; Conservative 105; Mismatches 263; Indels 172; Gaps 33;

QY 1 SPSDEKNDYINNQSI-----NTDNNQIIKKEETNNYDGIEKRESDRTSEST--- 48
 Db 180 NTDNNSDVANQNEPVAQNDKATSNEDVYASDYKDDGHSDNADDLTDQNESETLND 239

```

OY 49 -NVDENEATFLOKTPDNTH-----LFE--EEVKESSSVESSEN---SSIDTAQOPSH 94
Db 240 NAVSSNEDVASSDVAKODGTHSDDNASDDLTIDQNESVAQNDKAETSNEEDVASSDVAKODPTH 299
OY 95 TTIN-----REESVQTSNVEDSHVSDFANSKIKESNTEG-----KEENTIEQPN- 140
Db 300 SDANASDVADQNESETQNDNKNESNEDDVASSDVNODDTHSDANASDDVAKQDNESQND 359
OY 141 -----KVEDSTTS--QPSGYNIDEKISNODELNLPINEYENKARPLSTT 185
Db 360 KAETSNEEDVASSDVAKODDTHSDANASDIADQNESVAQNDKA-----ETSNEEDVASS 411
OY 186 SAQBSIKRYVYNQLAEQGSVNHILKYT---DOSTIEGYDSS---EGVYK-AHDAENLI 238
Db 412 DKODDTHSDANASDIADQNESATODDKATSKEDDVASSNDKQDNKAVSNIIKEASTAENKY 471
OY 239 YDVTF--EVDKVKSGDTM--TVDIDKNTVPSDLTDSFTPIKIKONGSGEIIATGYDNKN 294
Db 472 QPATEFSKVTPKLKVATTSANTAVATRSATVKEATRALPKY-----S 515
OY 295 KOITYTFTDYYDKYENIKA--HLKLSYIDSKVPNNNTKLDVEYKTAISSVNTTIVE 351
Db 516 PKVNSSINNYIRK--NNEKAPVYEDDIANYLPQYNYRYGKPEGIVMHDPA--NDNSTITGE 572
OY 352 YQRENERNTANLOSMFTNIDTKNHTVEQTIYINPLRYSAKETNWNISGNGDEGSTIIDS 411
Db 573 INYKKNNTY---SAFVAVYVDGDRIIETANTDYLANGA-----GPOANDRF 615
OY 412 TIIRYKVGDNQNL.PDSNRIVDYSEYEDVTNDYAOLGNNNDVANINFGNIDSPYLIKVIS 471
Db 616 IHELVHRTHDYDSFARS--INNADYA-ATNLQYGYLVPD--SAEYDGVFTWTHQAVS 669
OY 472 KY-----DPN-----KDDYTTIQQVTVMQTTINEYTGFEFTASYDNNTIAFST- 513
Db 670 NYLGSDHSDPHGLAANVSYDELYDLIEKYLITG-----QAAWGTSSSGSTG 721
OY 514 -----SSGQGO-GDLPEPK--TYKIGDYVWEDVDKQDIONTNDNEKPLSNVLVTLTYPDGT 566
Db 722 GTGSGTSGNGTTPPSKSGTVKV-----TENNGVRINSK-----NDGLYTVVYDQKG 770
OY 567 SKSVRTDED-----GKYO 580
Db 771 KTDPRVNOTLKVTKSATIGKEQY 793

```

Search completed: March 14, 2003, 13:01:24
 Job time : 16.8208 secs

GenCore version 5.1.4-F5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 7.64875 Seconds

(without alignments)
3155.972 Million cell updates/sec

Title: US-09-147-405b-13

Perfect score: 3013

Sequence: 1 SSDEKNDYNNNQSINTDD.....PDGTSKSVRTDEGKQYDFDG 582

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300.5	10.0	1018	FNBA_STAU	P14738 staphylococ
2	203	6.7	537	ARP_PLAFA	P04931 plasmodium
3	202	6.7	1093	PIA_KDICI	P54677 dictyosteli
4	196.5	6.5	1139	HML1_MTCGE	Q09413 mycoplasma
5	192	6.4	918	YMBJ_CAEEL	P34487 caenorhabdi
6	189.5	6.3	1664	INTL_CANAL	P33705 candida alb
7	188.5	6.3	578	LIPA_MYCPU	Q50274 mycoplasma
8	188.5	6.3	1435	EBAL_PLAFC	P19214 plasmodium
9	186.5	6.2	3110	LMAR2_HUMAN	P24043 homo sapien
10	184.5	6.1	1744	TANA_XENLA	Q01550 xenopus lae
11	184	6.1	817	YG4A_YEAST	P46948 saccharomyc
12	184	6.1	1251	RBP2_PLAVB	Q00799 plasmodium
13	183.5	6.1	2009	SEC7_YEAST	P11075 saccharomyc
14	183	6.1	2452	RPBL_PLAFD	P14248 plasmodium
15	182.5	6.1	853	YCGI_YEAST	P25588 saccharomyc
16	182.5	6.1	1381	YBE7_YEAST	P34236 saccharomyc
17	182.5	6.1	1658	YME7_YEAST	Q03661 saccharomyc
18	182	6.0	1419	ALAI_CANAL	Q13368 candida alb
19	182	6.0	1460	N159_YEAST	P04279 saccharomyc
20	181.5	6.0	1093	YKDS_CAEEL	Q03563 caenorhabdi
21	179.5	6.0	1165	YNF4_YEAST	P33950 saccharomyc
22	179	5.9	1233	DSPF_HUMAN	Q91244 homo sapien
23	178	5.9	960	YMX6_YEAST	Q04279 saccharomyc
24	178	5.9	1875	YME1_YEAST	Q02455 saccharomyc
25	177	5.9	2660	YEEJ_ECO57	Q08877 escherichia
26	176	5.8	1196	BXCN_CIOBO	P46081 clostridium
27	175.5	5.8	1183	CNA_STAU	Q53654 staphylococ
28	175.5	5.8	1637	MSPF_STAU	P80544 staphylococ
29	175	5.8	1147	CGAI_HELPO	P80200 heliobacte
30	174	5.8	2022	ANT1_ONCVO	P21249 onchocerca
31	173.5	5.8	967	Y5G0_CIOAB	P33747 clostridium
32	173	5.7	1310	YB35_SCHPO	O14340 schizosacch
33	172	5.7	682	NISP_LACIA	Q07596 lactococcus

34	170.5	5.7	688	1	LIP_STAEP	Q02510 staphylococ
35	170.5	5.7	970	1	PSU1_YEAST	P33550 saccharomyc
36	170.5	5.7	1260	1	ALSI_CANAL	P46580 candida alb
37	170.5	5.7	1435	1	LTEL_YEAST	P07866 saccharomyc
38	170.5	5.7	1636	1	BUD3_YEAST	P25558 saccharomyc
39	170	5.6	1803	1	YJL3_YEAST	P47024 saccharomyc
40	169.5	5.6	571	1	TACY_STRPY	P21131 streptococc
41	169.5	5.6	837	1	CSG_HAIWO	P25062 halobacteri
42	169.5	5.6	1020	1	BCA_STRAG	Q02182 streptococ
43	169.5	5.6	2859	1	RBP1_PLAVB	Q00798 plasmodium
44	169	5.6	731	1	BAF1_YEAST	P14164 saccharomyc
45	169	5.6	954	1	BIR1_YEAST	P47134 saccharomyc

ALIGNMENTS

RESULT 1	FNBA_STAU	STANDARD	PRT	1018 AA.
ID	FNBA_STAU			
AC	P14738			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Fibronectin-binding protein precursor (FNBP).			
GN	FNBA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_Taxid=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NCITC 8325-4;			
RX	MEDLINE=89098998; Pubmed=2521391;			
RA	Signaes C., Raucel G., Joensson K., Lindgren P.-E.,			
RA	Anantharamiah G.M., Hoeck M., Lindberg M.;			
RT	"Nucleotide sequence of the gene for a fibronectin-binding protein			
RT	from Staphylococcus aureus: use of this peptide sequence in the			
RT	synthesis of biologically active peptides.;			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).			
CC	-1- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN			
CC	PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE			
CC	WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO			
CC	THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S. AUREUS,			
CC	THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE			
CC	INVASION.			
CC	-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib.ch).			
CC	-----			
DR	EMBL; J04151; AAA26632.1; -			
DR	InterPro; IPR004237; Fn_bind.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF02986; Fn_bind; 1.			
DR	TIGRfams; TIGR01167; LpxTG_anchor; 1.			
DR	TIGRfams; TIGR01168; YSIRK_signal; 1.			
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.			
DR	Cell wall; Peptidoglycan-anchor; Repeat; Signal.			
FT	SIGNAL	1	36	
FT	CHAIN	37	985	FIBRONECTIN-BINDING PROTEIN.
FT	PROPEP	986	1018	REMOVED BY SORTASE (POTENTIAL).
FT	REPEAT	545	574	B-1.
FT	REPEAT	575	604	B-2.
FT	DOMAIN	745	878	4 X APPROXIMATE TANDEN REPEATS,
FT				FIBRONECTIN-BINDING DOMAIN.
FT	REPEAT	745	782	D-1.

FT REPEAT 783 820 D-2.
 FT REPEAT 821 859 D-3.
 FT REPEAT 860 878 D-4 (INCOMPLETE).
 FT DOMAIN 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
 FT REPEAT 879 892 WR1.
 FT REPEAT 893 906 WR1.
 FT REPEAT 907 920 WR3.
 FT REPEAT 921 934 WR4.
 FT REPEAT 935 948 WR5.
 FT SITE 982 986 LPYTG SORTING SIGNAL (POTENTIAL).
 FT MOD_RES 985 985 AMIDE-LINKED TO CELL WALL (POTENTIAL).
 SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;

Query Match 10.0%; Score 300.5; DB 1; Length 1018;
 Best local Similarity 21.2%; Pred. No. 6.8e-06;
 Matches 123; Conservative 118; Mismatches 243; Indels 97; Gaps 25;

35 GIKRSEDREESTTNDENATFLQKTPDQNT---THLTFEEVKESSESSSSIDTAQ 90
 30 GQDEKAASQKTTVEEN-----GNSATDNKTSQTOTATVNHLEETSYNATV--TE 82
 QY 91 QPSHTT-INREE---SVQTSNDVEDSHVDFANSKIKESSTESGKEENTTEOPNKYEDS 146
 83 QPBNATQVTEEARPKAVQAQTAQNPANIEVEKREYKE-----EAKPOYKETT 130
 QY 147 TTSQPSG---YTNIDEKISODELLNPINEYENKARPLSTSNOPSIRKVT-----V 196
 131 QSDNSGQDQROVDLTPKATQNOVAETQV---EVAQPRASESKPRVTSADVAEAKEA 186
 QY 197 NOLAEGGSVNHLLIKYTDQSTEGYDSEGVKKAHAEMLIDYFVEVDKYSQDGTMT 256
 187 SNAKVEGTIVTSKYVEIGSI-BGHNNINKV-EPHAGQAVALKYKLFENGHLGGYFD 244
 QY 257 VDIKNTVPSDLTDSFTPIKIDNS-----GEIATGYDNKNQOITTYTETDYDKYENI 311
 245 FTLSNNVHTGVSATKRVPEIKGSVYMATGEVLEGS-----KITYTETNOIEDKVYD 297
 Db 312 KAHKLTSLYDKSKVPNNNTKLDVEYKTAASSVN-----KTIYEVORPNDENTANLQSM 366
 298 TAELEMLFTDPTVQVQNGN-----OTITSLNEQTSKELDYKYGIGNYVYANGS 351
 QY 367 FTNIDFNHTVEQTYINPLRYSAKETNVNISGDEGFTIIDSTLIKKYK-VGDQNL 425
 352 IEFNKNNNPFSHVAFKFP--NNGKTSVYVTGTLMKGSNONGNQPVRIRFEILGNEDI 409
 Y 426 PDSNRIV---DYSEYEDVYNDYDQQL--NNNDVINFINGIDSPLYIKVYSKYDPKND- 478
 410 AKS--VYANTTDTSKFEVTSNMSGNLNONGSYSLNIELOKTYVYVHDEGLNGTDE 467
 QY 479 -DITTIQOYVTMOTTINIEYTGERTASYDNTIAFSTSSGQCGDLP-----EKTYSIGD 532
 Db 468 VDEFT--OMVGHPEQLYKYUDRGYTLTWMDGLVLSNRKANGNEKNPILIONNKEFYK--- 523
 QY 533 YVMEVDKQGIQNTNDNEKPLSNLVLTLPDGTSKSVRTD 573
 Db 524 ----EDTIKETLIGQYDKN-----LVTTVEEYDSSTLIDID 555

RESULT 2
 ARP_PLAFA STANDARD; PRT; 537 AA.
 AC P04931;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Asparagine-rich protein (AG319) (ARP) (Fragment).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66176787; PubMed=2421257;
 RA Stahl H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,

RA Brown G.V., Anders R.F., Kemp D.J.:
 RT "An asparagine-rich protein from blood stages of Plasmodium
 RL falciparum shares determinants with sporozoites.";
 CC Nucleic Acids Res. 14:3089-3102(1986).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-stb.ch).

CC EMBL; M24328; AAA29491.1; -
 DR PIR; A23770; A23770.
 KW Malaria.
 FT NON_TER 1 537
 SQ SEQUENCE 537 AA; 63000 MW; A6D76B15318C239 CRC64;

Query Match 6.7%; Score 203; DB 1; Length 537;
 Best local Similarity 17.8%; Pred. No. 0.037;
 Matches 93; Conservative 106; Mismatches 212; Indels 112; Gaps 19;

QY 11 NNNOSTDNDNOIIEKEETNNYDGIKRESDREESTTNDENATFLQKTPDQNTLITE 70
 Db 4 NNNNNNNNDGNN--INVQNTNEKDNKNKNNFKQYNNNKFENNNNSNTMHSRNSVE 61
 QY 71 EEVKESSESSSSSID--TAQPSHTT-----NREESVQ---SDNEDSHVDFAN 119
 62 EHLR-ANSLMNNNSNTNNYNOOTRSPFMENENENKRYHGGANNNHFKRYKYYNNN 120
 QY 120 SKIKESNTESGKEENTTEOPNKYKEDSTSPSGYTNIDEKISODELLNP-PINEYENK 178
 121 SSMK--NTDNNKTDTSYMAKGTINNDN-----NMODYLRINIINNIEYKGS 163
 QY 179 AREPLTSSAOPSIKRYVNDOLAEGGSVNHLLIKYTDQSTIEGVDSEGVKADANLI 238
 164 AKKKEFTYN-----YMKRNLKFTQNNNDNMNINEDNNNNNNNNNGVFSNQNNNMN 217
 QY 239 YDVFEDVDKSGDGTVTVIDKNTVPSDLTDSFTPIKIDNSGEIATGYDNKNQOIT 298
 218 RNNNSINKRLN-----NNNNNNNNNNKKSQKQNSNNNPFYNNYQARKKSMN 268
 QY 299 YTFDYVYKXENIKAHKLTSLYDKSKVPNNNTKLDVEYKTAASSVN----- 345
 269 NNN 328
 QY 346 -KTIYEVORPNDENTANLQSM-----FTNIDFNKNTV-----EDTIYIN----- 384
 Db 329 VKASMMYNN-NESNTANPQOMNEQTNNDMKRENNNNNNYGYDDTVVHNNNTPTSDP 387
 QY 385 -----PLRYSAKETVNI-----SCNGD-----EGSTIIDSTLIKKYK-----VG 420
 Db 388 FSRAYGVYNNYNNNNNNNSGNNMKKRENNENKRVADNNSLNNKNNNNNNIN 447
 QY 421 DNQNLPSNRIVDYSEY-----EDVTNDYDQQLGNNNDVIN 457
 Db 448 MNEISNNNNNTLNNNNEYNQNNNEDDDDDMGELGEDKXYIDIN 490

RESULT 3
 PI4K_DICDI STANDARD; PRT; 1093 AA.
 AC P54677;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphatidylinositol 4-kinase (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-
 DE kinase) (PI4K-alpha).
 GN PI4K OR PI4K.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.

NCBI_TaxID=44689;
 OX 11
 RN SEQUENCE FROM N.A.
 RP STRAIN-AX3;
 RC MEDLINE=96009592; PubMed=7565716;
 RX Zhou K., Takegawa K., Emr S.D., Firtel R.A.;
 RA "A phosphatidylinositol (PI) kinase gene family in Dictyostellum
 RT discoidum: biological roles of putative mammalian p110 and yeast
 RT Vps34p PI 3-kinase homologs during growth and development.";
 RL Mol. Cell. Biol. 15:5645-5656(1995).
 CC -1- FUNCTION: ACTS ON PHOSPHATIDYLINOSITOL (PI) IN THE FIRST
 CC COMMITTED STEP IN THE PRODUCTION OF THE SECOND MESSENGER
 CC INOSITOL-1,4,5-TRISPHOSPHATE.
 CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1L-myo-inositol = ADP +
 CC 1-phosphatidyl-1D-myo-inositol 4-phosphate.
 CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U23479; AAA85725.1; -.
 CC DictyDb: DD01102; P1KD.
 DR InterPro: IPR000403; P13_P14_Kinase.
 DR Pfam: PF00454; P13_P14_Kinase. 1.
 DR SMART: SM00146; P13Kc. 1.
 DR PROSITE: PS00915; P13_4_KINASE.1; 1.
 DR PROSITE: PS00916; P13_4_KINASE.2; 1.
 DR PROSITE: PS50290; P13_4_KINASE.3; 1.
 DR Transferase: Kinase; Multigene family.
 FT DOMAIN 146 156 ASP-RICH.
 FT 184 200 POLY-ASN.
 FT 203 206 POLY-ASN.
 FT 227 233 POLY-ASN.
 FT 277 315 POLY-ASN.
 FT 441 445 POLY-ASP.
 FT 445 466 POLY-THR.
 FT 494 501 POLY-GLY.
 FT 501 501 POLY-THR.
 FT 690 696 POLY-THR.
 FT 751 754 POLY-GLN.
 FT 761 772 POLY-THR.
 FT 775 785 POLY-THR.
 FT 833 1093 P13K/P14K.
 SQ SEQUENCE 1093 AA; 123017 MW; 8762BC78355AA635 CRC64;
 Query Match 6.7%; Score 202; DB 1; Length 1093;
 Best Local Similarity 20.0%; Pred. No. 0.093;
 Matches 116; Conservative 100; Mismatches 223; Indels 142; Gaps 24;

266 SDLTDSFTIRKIDNSG---EIIATGYDNKKNOITFTFDYVDKYEINIKAHKLTSYID 322
 Db 469 NHLSTKTSVGINSSNSTPININSAGAGAGGGGGINH-----GYDI-----SYLD 515
 OY 323 KSKVPNNNTKL---DVEYKTA-----LSYVKRTIVE---VORPENTANIQS 365
 Db 516 KCKTPPESKLSHDPEFELSRCYLDILSFIOKLHISILPLPIDLRQAKLKH 575
 OY 366 MFTNIDKRNHTVEQTIYINPLRYSAKETNVNISGNGEGSTIIDSTTIKRYKVDNQL 425
 Db 576 ---EISLILNINLPLGLIV-PLMQSS-----NHHGVVRI-----P 605
 OY 426 PDSNRIDYSG-----YEDVTDNDYAOLGNNNDVNINFGNIDSPYILIKVSKYDPKDD 479
 Db 606 PEEVKILNSRRRVPFLLVLEIESEHLSN-----LFEVVSF-----LQ 647
 OY 480 YFTIQGTFTVMTQTYNEYTGE---FRTASVDNTIAFSTSS 515
 Db 648 YTGNSALKKDDIRKRYSEKFKSFINSINSTSNSSDS 688
 RESULT 4
 HMW1_MYCGE STANDARD; PRT; 1139 AA.
 AC 049413; 049365;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome high molecular weight protein 1 (Cytochrome accessory
 DE protein 1).
 OS HMW1 OR MG312.
 GN Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischman J.L., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudak D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Luster J.C.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RT Science 270:397-403(1995).
 RN 12
 RP SEQUENCE OF 721-847 FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing.";
 RT J. Bacteriol. 175:7918-7930(1993).
 RL 1- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
 CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
 CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW
 CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHESIN PROTEINS
 CC IN THE MYCOPLASMAL MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: LOCALIZES SPECIFICALLY TO THE ATTACHMENT
 CC MEMBRANE (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U39712; AAC71534.1; -.
 DR EMBL: U02261; AAD12527.1; -.

DR TIGR: MG312: -
 KW Cytochrome: Structural protein: Complete proteome
 SQ SEQUENCE 1139 AA; 13051 MW; 0011D3288C3D0836 CRC64;

Query Match 6.5%; Score 196.5; DB 1; Length 1139;
 Best Local Similarity 18.0%; Pred. No. 0.17;
 Matches 150; Conservative 116; Mismatches 276; Indels 293; Gaps 34;

```

QY 1 SSDEKADVNNNOSTITDNNQI--IKKE-----ETNNYDGIKRSDDRT 44
DB 343 SSSFEETPTVEOPFQVNVSEVNDQKPEITKEPVLESFKNQDVVETSDLSNENLSENN 402
QY 45 ESTNNVENATFLQ-----KTPQDNTH-----LTSEVK 74
DB 403 KDATNNDLSNSETQLNSSETASDVHYESKSEPIHDYKFGSOLSQSNNSNLSSEPVK 462
QY 75 ESSSV-----ESSNSSID-----TAQPSHTTIRRE----- 101
DB 463 FNSETAPDAHFESQSEVVDQVYIQNEELKPTLDPPSSDDYAKQPTDENYGFNDLP 522
QY 102 -SVQTSNVEDSHVDFANSKIKESNTSGKEENTIEOPNKVKEDSTTSOPS----- 152
DB 523 PEVKQPSVVDQPSDDHFAKQPESTDSYSPFSDLP-----TLDPQSLDDHYQYN 575
QY 153 -----GTNDEKISNDEL-----LNLP 171
DB 576 FDHHEELKPAVEQNNYQVGFDOQANLNDNEELQPTAEKVTTFESKOAVVDSYOLP 635
QY 172 IN-EYENKAPPLSTTSQSIKRVYVNLAAEQSNVNHLLK--VTQSTIEGYDSEGV 228
DB 636 IDTQDQGTFFSSSFQPTVEQFD-----QVNSEVNDQKPEITKEPVLESFKNQDV 689
QY 229 IKAHDAENLYDVTEVEDKV-----KSGDTMTVDIDKN 262
DB 690 VETSNYNNQKQEDIQSDNKITITTKKSSQPIPTLPISFVSNIEKPVETTLADKNES 749
QY 263 -----TVPSDLTDSFTIPK-----IKNSGEIATG--TYDNKNGKQITTFYVD 306
DB 750 QQEQITNTSTEDSKTLAKTLVOLQOINLNOSTIVTSVSLDKDDQDITNTVASED 809
QY 307 KYENIKAKLITSYIDSKVPNNNTKLDVEKKTALSSVNK-----TTVEYQRPENRTA 361
DB 810 QQRPIEVFAKKEVEHESTQN--KQSVEDKSELDFNKKSKDLKIISELKRELPLTI 867
QY 362 NLSMFTNIDTKNHTVQGTIYNPL-----RYSAKETVNI 397
DB 868 NFDATFQNMNYQMSVKOSFIHLNDFVNTYKQISERYLLIKKELOSELSRLIDQENMLN 927
QY 398 SGNDSESTIIDSTIIR-----YKYGDN-QNLPDSKRIYD----- 434
DB 928 QFNNAKMLTTLQKEEMRSLASDAFAIAYKPSNSYEOLQSGEIMRHQORAITENKIES 987
QY 435 -----SEYEDVTNDYQALGNN--NDVNINFGNIDSPYIIKIVISKYD-----PNK 477
DB 988 IQGSLKGLKTYVNSCCETIMNNIKKLDNTLFAKKEKDPPL--LSNDSVYDNGLVEPQ 1045
QY 478 --DDYTIQGTVMQTTINEYTGFEPTASVYDNTIAFSTSSQGGGLPEKTYKIGDYV 535
DB 1046 LMDL--LIDSNFTPDNISNEQLDPIYEMMDRNIIDFEFEGFN-----NDFV- 1089
QY 536 EDVQ--KDIGQNNDNEMKPLSNVLYLTYPDTSK--SVRDED-----GKYQFD 581
DB 1090 -DIDAKVMSAFSVNDLLETLY-----PDRTSNFSLLDEDLFESSGGFSLD 1138

```

RESULT 5
 YMB_CAEEL STANDARD; PRT; 918 AA.
 AC P34487;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 96.6 kDa protein f59B2.12 in chromosome III.

GN F59B2.12
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bliscol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favallo A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,
 RA Wohldman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z11505; CAAT7581.1; -
 DR PIR: S31132; S31132.
 DR WormPep: F59B2.12; CE01024.
 KW Hypothetical protein.
 SQ SEQUENCE 918 AA; 96560 MW; E464FD86B1495DE CRC64;

Query Match 6.4%; Score 192; DB 1; Length 918;
 Best Local Similarity 21.2%; Pred. No. 0.2;
 Matches 138; Conservative 90; Mismatches 278; Indels 146; Gaps 26;

```

QY 4 EKKDVNNNOSTITDNNQIKEETNNYD-----GIKRSSEDRK-----STN 49
DB 113 ESKVDEANENKIKRSAGS-VIERGSKHSSDASSYGLKSSKYADKNGTILMSYTN 171
QY 50 VDENEA-----TFLOKTPQDNTHL-----FEEVKESSVSSSSIDTAQOP 92
DB 172 KINNQSALDEGNEFYVNOQADGTFLRNNTGKKNIDESHVNLVDENAGMSTG-ADGT 230
QY 93 SHTTINREESVQTSN-VEDSHVDFANSKIKESNTSGKEENTIEOPNKVKEDSTTSOP 151
DB 231 SHNTINRKGVSVDGHNMAADAH-----SNFESLDAQNGKKSSQNYSKKAASAG- 278
QY 152 SGYINIDEKISNOBELNLNPINEYENKARPLSTTS-----AQPSIKRTVN-----OLA 200
DB 279 --SNAQFE-SNLESLKNAADGTSMSNSTGNENNTSYDKATAEVWSKKNVADGTSMEA 334
QY 201 AEQSNVNHLLKYVDOSTTEGYDSEGVYKADHAENLYDVTEVEDQVKKSGDTMTVID 260
DB 335 SHAGSNSKINSASGQSSDLSMGPNG-IKSHSTSNKTDN--YALDENQAGSISSEIG 391
QY 261 K-----NTVPSDLTDSFTIPIKINDNGEIIATGYDNKKKQIYTT 300
DB 392 KNGCRSLNESSIESGRKAESRNNTAADTLSDVANGVTSSSHSKSASSTSDENHNKTHA 451
QY 301 FTDYVDYENIKALKLITSYIDSKVPNNNTKLDVEYKTA--LSSVN-----KTTTV 350
DB 452 LQASVDEHGNNKNSIDGSRNKKRTGEGNSEMSASIKNAGDTISQVAVKNDTRNNYEA 511
QY 351 EYQRPENRTPANLQSMFTNIDTKNHTVQGTIYNPLRYSAKETVNIISGNDSESTIID 410
DB 512 EKSALERNHEKNSDGTFRD-----ESKGSNRYN-RIDGGSNLAIVG 551

```

QY 411 STIKYKVDNONLPDSNRITYSEEDYNDYQALGNNDVNGINDSPIIKYI 470
 FT CARBOHYD 817 817 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 838 838 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 841 841 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Db 552 SYSGKGVSSNETIASSN-AFNTSDAES--NQFDHLQKRTANGELITHAK-DS-----KQV 604
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1082 1082 N-LINKED (GLCNAC. . .) (POTENTIAL).
 QY 471 SKDPPKNDYTTTIOQYTMQTTINETYGERFATYDNTIAFTSTSSGQSGDLPEPKYKI 530
 FT CARBOHYD 1100 1100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1113 1113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Db 605 AASANAASSLDTSMSAIDAGNKVKDTSSQAADSHD---AISASS----- 646
 FT CARBOHYD 1200 1200 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1200 1200 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1571 1571 N-LINKED (GLCNAC. . .) (POTENTIAL).
 QY 531 GDVMEVDVDKGIQNTNDNEKPLSNVLVTLLYPPDGTSKSVYEDDGKVFQPDG 582
 FT CARBOHYD 1593 1593 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 106 POLY-GLN.
 FT DOMAIN 283 286 POLY-ASN.
 FT DOMAIN 1283 1292 POLY-HIS.
 FT DOMAIN 1651 1660 POLY-GLN.
 Db 647 1-----DVDAKIVKHA--DRSEISISDSSNOTASEHNDSSKQSEHEKRONADG 691
 SQ SEQUENCE 1664 AA: 187859 MW: 8426000.000000 CRC64:
 RESULT 6
 INTL CANAL STANDARD: PRT: 1664 AA.
 ID INTL CANAL STANDARD: PRT: 1664 AA.
 AC P53705;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Integrin alpha chain-like protein (Alpha-INT1).
 GN INT1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10261;
 RX MEDLINE=96133936; PubMed=8552638;
 RA Gale C., Finkel D., Tao N., Melnik M., McCellan M., Olson J.,
 RA Kendrick K., Hostetter M.;
 RT Cloning and expression of a gene encoding an integrin-like protein
 RT in Candida albicans.
 RL Proc. Natl. Acad. Sci. U.S.A. 93:357-361(1996).
 CC - FUNCTION: COULD PLAY A ROLE IN ADHESION AND IN STE12-INDEPENDENT
 CC MORPHOGENESIS.
 CC - SUBCELLULAR LOCATION: CELL-SURFACE OF THE PLASMA MEMBRANE.
 CC - SIMILARITY: SOME, WITH INTEGRIN ALPHA CHAINS.
 CC - SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U35070; AAA96019.1;
 CC Interpro: IPR001849; PH.
 JR Pfam: PF00169; PH: 1.
 DR SMART: SMO0223; PH: 1.
 DR PROSITE: PS50003; PH_DOMAIN: 1.
 DR Cell adhesion; glycoprotein.
 FT SITE 1157 1151 PH.
 FT DOMAIN 1527 1636 CELL ATTACHMENT SITE (POTENTIAL).
 FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 583 583 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 817 817 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 838 838 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 841 841 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Db 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1082 1082 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1100 1100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1113 1113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1200 1200 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1571 1571 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1593 1593 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 106 POLY-GLN.
 FT DOMAIN 283 286 POLY-ASN.
 FT DOMAIN 1283 1292 POLY-HIS.
 FT DOMAIN 1651 1660 POLY-GLN.
 SQ SEQUENCE 1664 AA: 187859 MW: 8426000.000000 CRC64:
 Query Match 6.38; Score 189.5; DB 1; Length 1664;
 Best Local Similarity 20.48; Pred. No. 0.51;
 Matches 131; Conservatively 105; Mismatches 236; Indels 171; Gaps 31;
 QY 8 DVNNNS--INTDDNQIILK-----EETNNVDGIEKRSDETESTT 48
 Db 138 DVNNHAPYINTSPKSIKKATPKASPCKVAFTVNPPIHIF--DNRYEEDDSQ 195
 QY 49 NVDENEATFLOKTPQDNTHTLEEVKESSSVSSSSIDTAQOPSHT-----INRE 101
 Db 136 KEDSVPEPLIQ-----HOKKDPQFVYSDPDNNAVPT-PLHTTKPTPAQLNNNN 247
 QY 102 SVQTSNVEDSHVD-----FANSIKI-----SNTSEKGEENTIEQPNKYKEDSTT 148
 Db 248 EV-----NSEPALDMKLRKRENSLSLDEYVNLVYLSPTNNNSKNVSDMSHQ--N 299
 QY 149 SQPSGYTNDIKISNODLLNLPINENKARPL-STSAOPSIKYVNVQLAAGSNV 207
 Db 300 LODASKNKTNNINHLSPALAPKNDIEN--PLNLSLNAISLSKSSSSSLOSLRND 356
 QY 208 NHLIKV--TQSTIEGYDSEGVYKADAENLIDYFEVDKVSQDTMTVIDKNTV 264
 Db 357 NRVLSEVSPKPKVNPGLSLNDGIKGFSD-----EVEVELL 392
 QY 265 PSDLT-DSFTIPKID-----NSGEIATGYDNKNOITTFDYDKYENIKAH----- 314
 Db 393 PROLSRDLKLETKEDAPENHNEFIDAKSTNKG-OLVYSSDHDLSFPRSYNHTBQS 451
 QY 315 -LKLTSYIDKSKVPPN-----NKLDEYKKTALSSVYNK 346
 Db 452 ILNLSASQSOISLNALEKQROTOEOTOAAPEEETSNDIKVQEKSNLEFYKV 511
 QY 347 TIVVEQRPNNRNTANLOSMTNIDTKNH-TVEQTIYINPLYSAKETNNVIG----- 399
 Db 512 TIKKEPVASATEIKAPKREFSRILRIKNEDEIAEPADIHPPKENANSHVEDTALLKA 571
 QY 400 -NGDEGSTIIDSTIYK-YKVDNONLPDSNRITYSEYEDYNDYQALGN--NDV 454
 Db 572 LNDESDPTONSTKMSIRFHIDSDMKLEDSND--GDRE--NDIDSRKESDILNDV 625
 QY 455 N-----INGNISPTIYIKISKYDPNKKDDYTTIOQYTMQTTINETYGERF-----T 502
 Db 626 SOTSIIIDKYGNSSSEITTKLA---PPRSNNNKKNSKSLIEDANNESLQQLLEVHT 682
 QY 503 ASYDNTIAFTSSSGOGDLPPEK---TYKIGDY-VVEDYDK 540
 Db 683 KEDDSTLANSSNLA-----PPEELTLVPEANDYSFNDYVK 719
 RESULT 7
 LIPA_MYCPU STANDARD: PRT: 578 AA.
 ID LIPA_MYCPU
 AC Q50274; Q98039;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lipoprotein A precursor.
GN LIPA OR MYPU_5300.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KD735-15;
RX MEDLINE=20245550; PubMed=10781561;
RA Shen X., Gumalak J., Yu H., French C.T., Zou N., Dybvig K.;
RT "Gene rearrangements in the *usa* locus of *Mycoplasma pulmonis*.";
RL J. Bacteriol. 182:2900-2908(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chandraud I., Hellig R., Ferris S., Barbe V., Samson D., Galisson F.,
RT Moszer I., Dybvig K., Wroblewski H., Viart A., Rocha E.P.C.,
RN Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT *Mycoplasma pulmonis*.";
RL Nucleic Acids Res. 29:2145-2153(2001).
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RC STRAIN-KD735-15;
RX MEDLINE=96414471; PubMed=8817492;
RA Bhugra B., Voelker L.L., Zou N., Yu H., Dybvig K.;
RT "Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven,
RT site-specific DNA inversions.";
RL Mol. Microbiol. 18:703-714(1995).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (potential).
CC -1- SIMILARITY: BELONGS TO THE PULMONIS LIPAB LIPOPROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U23947; AAB41030.2; -
DR EMBL: AL445565; CAC13703.1; -
DR Mypulist; MIPU_5300; -
DR y Lipoprotein; Membrane; Signal; Complete proteome.
CC
FT CHAIN 1 27 POTENTIAL.
FT LIPID 28 578 LIPOPROTEIN A.
FT LIPID 28 28 N-ACYL DIGLYCERIDE (POTENTIAL).
FT DOMAIN 82 165 ASN-RICH.
FT VARIANT 7 7 K -> R (IN STRAIN KD735-15).
FT VARIANT 41 41 N -> S (IN STRAIN KD735-15).
FT VARIANT 110 110 S -> D (IN STRAIN KD735-15).
FT VARIANT 127 127 S -> A (IN STRAIN KD735-15).
FT VARIANT 547 547 Y -> N (IN STRAIN KD735-15).
SQ SEQUENCE 578 AA; 66219 MW; B533ED3467005DB9 CRC64;
Query Match 6.3%; Score 188.5; DB 1; Length 578;
Best Local Similarity 22.0%; Pred. No. 0.16;
Matches 133; Conservative 81; Mismatches 257; Indels 133; Gaps 30;

OY 174 EYENKARPLSTTSAPSIKRTVYNQLAEOGSVNVHLIKV-----TDQSIT--EGYD 223
DB 194 NRQNK-----NEKKDDKKALQWQKLNESASILESSTYQTSLSLTFKBCMP 240
OY 224 -DSEGVKAKDAENLIYDVFVEVDKVGSGDTMTVDIDKNTVPSD--LTDSTFIPKIDN 280
DB 241 LGLEVLAK---LENLDSHEEKESIFKTTNGKVQVNLTLSSNLTSCKMKIKSFSDKTYSH 297
OY 281 SGEIATGYDNKKN---QITTYTVDYDKENIKAKHLK-TSYIDK---SKVNNNTKL 333
DB 298 SPTEETFEDEKTNKLEKOEERIEKIOKADIKYQ-----IKIKQNKYDKPLLSIKVALSNEDL 352
OY 334 DVEKTKLSSVKNKITVEYQRPNNRNTANLSMFT-----NIDPKN---TVQGT 380
DB 353 NLNFRD-LEIYVNSLKFSGSNLNDLSEKOKITFEKVSENOAKTKRAFKITTLTDLN 411
OY 381 IYIPLRYSAKETVWNISGNGDECSIIDD--STIIKYK--VGDNONLTPDS---NRI- 431
DB 412 VFEKTLKMSYK-TNKEVLDEFKNSALMDLOQLSITVFEKSLMHPYQLPRAKSKINTIN 470
OY 432 -----YDYSEYEDVTNDYDQALGNNDVNVNFGNIDSPYIIVISKYDPNKDDY 480
DB 471 LINDVASFOGYDYL-----DDF-----NGSAKLKF-----KLQRCGEQORDIT 508
OY 481 TTICQYTMQTTINTEYGEFTASYDNTIAFTSSGCGGGLPPE--KTYIGDYVWEDV 538
DB 509 FTINGFLKVLSDPLPYKGNLNSEPD-VKASSNGYLGQYTTAAEVFHNYSNGKSYWAT 567
OY 539 DKDG 542
DB 568 ANDG 571
RESULT 8
EBAL_PLAFC STANDARD; PRT; 1435 AA.
ID EBAL_PLAFC
AC P19214;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Erythrocyte-binding antigen EBA-175.
OS Plasmodium falciparum (Isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90377299; PubMed=2204835;
RA Slim B.K.L.;
RT "Sequence conservation of a functional domain of erythrocyte binding
RT antigen 175 in *Plasmodium falciparum*.";
RL Mol. Biochem. Parasitol. 41:293-296(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X52524; CAA36756.1; -
DR PIR: S11561; S11561.
KW Antigen.
FT DOMAIN 159 1104 ESSENTIAL FOR BINDING TO
FT VARIANT 1031 1031 ERYTHROCYTES.
FT VARIANT 1031 1031 E -> V (IN STRAINS FOR-3 AND ITG).
SQ SEQUENCE 1435 AA; 167389 MW; 32A4309021B1C3D6 CRC64;
Query Match 6.3%; Score 188.5; DB 1; Length 1435;
Best Local Similarity 20.5%; Pred. No. 0.47;
Matches 114; Conservative 102; Mismatches 199; Indels 141; Gaps 28;
OY 1 SDEEKNVNNQSIINTDNNQIIKKEETNNYDQ-----IEKRSDETESTINVDEN 53

```

Db      875 NSDSEEEVYVNH-----ISKSPINNGDSSGSAIVSSSSSSNGSLDDDRN 923
Qy      54 EATFLQRTPODNTLTLTEEV--KSSSVSSSSSIDTAQPSHTTINEESVQTSNVED 111
Db      924 GDFEVRTPQDPTANT-----EDVIKRENAKDEKDEKGADEERHSTSESLSPEEKMLTDNCGG 979
Qy      112 SHVSDFAANSKIKE--SNTESGKEEITIQPKVKE-DSTTSQPSQTYIDE----- 159
Db      980 NSLN---HEEVKEHTNSNDVQSGGIVNMVVEKELKDTLENPS--SSIDEKKAHELSE 1034
Qy      160 HINODELNLPINEYENKARPLSTSAQPSIKRVTVNOLAEGSGVMNLIKVTDOSI 218
Db      1035 PNLSSDDMSKTP-----GPLDNTS-EETTERISNNEKAVNE-----REDERTL 1077
Qy      219 TEGYDSEGVYKIAH-----DAENLIYDVTEFVDDKVKSGDTMVDIDKNTPSDLTDSF 272
Db      1078 TKYEDI--VLKSHMNRSDGE--LVD-----ENSLDSTYNDESEDAEAKMGND 1124
Qy      273 TIPKIKNSGIIATGYTDKNKKQITTYFDYVDKYEINIKAKLKTSTIIDSQVFNNTK 332
Db      1125 T-SEMSHNSOHISSDOCKNMKRYGDLGTHVON---EISVPVTGEIDEKLRSEKSEK 1179
Qy      333 LDVEYKTLSSVNTTITVEYQRPENRFRANLOSM--FTNIDTKNTVEQITVINPLRSA 390
Db      1180 IHKADEERLSTHD---IHKINP-EDRNSNTLHLKDINRENEERHLTMQNTINISQERDLO 1234
Qy      391 K---ETVNNISGNGDEGSTIIDSTIIVYKVGDNQNLPSNRIVYSEVEDVTDDYAO 447
Db      1235 KHGHTIMNNLHGDC-----VSEKQINSHSHGNQD-----R 1266
Qy      448 LGNNNDV-----NINSGNIDSPYIKVSKYD-----PNKDYTTIQOTVMQTTI 493
Db      1267 GGNNGVNLNMRSSNNNNENNIPSRY-----NLYDKKLDDLLENRNDSTT-KELIKKLAEI 1320
Qy      494 NEYGEFTASYDNTI 509
Db      1321 NKCENISVKYCDHMT 1336

RESULT 9
LMA2 HUMAN STANDARD; PRT; 3110 AA.
AC P24043; Q14736; Q93022;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy
chain).
DS LAMA2 OR LAMM.
NC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94124633; PubMed=8294519;
RA Vuolteenaho R., Nissinen M., Sainio K., Byers E., Eddy R.,
RT "Structure of the human laminin alpha2-chain gene (LAMA2), which is
RL affected in congenital muscular dystrophy."
RN J. Biol. Chem. 271:27664-27669(1996).
RP [3]
RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.

```

```

RC TISSUE=Placenta; PubMed=2185464;
RX MEDLINE=90238994;
RA Ethig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.;
RT "Merosin, a tissue-specific basement membrane protein, is a
RL laminin-like protein."
RN Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268(1990).
RP [4]
RP VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.
RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
RA Marzluft G.A., Amato A.A., Mendell J.R.;
RT "Novel single base polymorphisms and rare sequence variants in
RL the laminin 2-chain coding region detected by RNA/SSCP analysis."
RN Hum. Mutat. 13:174-174(1999).
RP [5]
RP ERRATUM.
RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
RA Marzluft G.A., Amato A.A., Mendell J.R.;
RL Hum. Mutat. 13:340-340(1999).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end.
CC -1- THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
CC 4 (S-MEROSIN).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: PLACENTA, STRIATED MUSCLE, PERIPHERAL NERVE,
CC CARDIAC MUSCLE, PANCREAS, LUNG, SPLEEN, KIDNEY, ADRENAL GLAND,
CC SKIN, TESTIS, MENINGES, CHOROID PLEXUS, AND SOME OTHER REGIONS OF
CC THE BRAIN; NOT IN LIVER, THYMUS AND BONE.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -1- DISEASE: DEFECTS IN LAMA2 ARE THE CAUSE OF MEROSIN-DEFICIENT
CC CONGENITAL MUSCULAR DYSTROPHY (MCMD).
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 226553; CA81394.1; -
DR EMBL: 066796; AAB18388.1; -
DR EMBL: 066733; AAB18388.1; JOINED.
DR EMBL: 066734; AAB18388.1; JOINED.
DR EMBL: 066735; AAB18388.1; JOINED.
DR EMBL: 066736; AAB18388.1; JOINED.
DR EMBL: 066737; AAB18388.1; JOINED.
DR EMBL: 066738; AAB18388.1; JOINED.
DR EMBL: 066739; AAB18388.1; JOINED.
DR EMBL: 066740; AAB18388.1; JOINED.
DR EMBL: 066741; AAB18388.1; JOINED.
DR EMBL: 066742; AAB18388.1; JOINED.
DR EMBL: 066743; AAB18388.1; JOINED.
DR EMBL: 066745; AAB18388.1; JOINED.
DR EMBL: 066746; AAB18388.1; JOINED.
DR EMBL: 066747; AAB18388.1; JOINED.
DR EMBL: 066748; AAB18388.1; JOINED.
DR EMBL: 066749; AAB18388.1; JOINED.
DR EMBL: 066750; AAB18388.1; JOINED.
DR EMBL: 066751; AAB18388.1; JOINED.

```

DR EMBL; 066752; AAB18388.1; JOINED.
 DR EMBL; 066753; AAB18388.1; JOINED.
 DR EMBL; 066754; AAB18388.1; JOINED.
 DR EMBL; 066755; AAB18388.1; JOINED.
 DR EMBL; 066756; AAB18388.1; JOINED.
 DR EMBL; 066757; AAB18388.1; JOINED.
 DR EMBL; 066758; AAB18388.1; JOINED.
 DR EMBL; 066759; AAB18388.1; JOINED.
 DR EMBL; 066760; AAB18388.1; JOINED.
 DR EMBL; 066761; AAB18388.1; JOINED.
 DR EMBL; 066762; AAB18388.1; JOINED.
 DR EMBL; 066763; AAB18388.1; JOINED.
 DR EMBL; 066764; AAB18388.1; JOINED.
 DR EMBL; 066765; AAB18388.1; JOINED.
 DR EMBL; 066766; AAB18388.1; JOINED.
 DR EMBL; 066767; AAB18388.1; JOINED.
 DR EMBL; 066768; AAB18388.1; JOINED.
 DR EMBL; 066769; AAB18388.1; JOINED.
 DR EMBL; 066770; AAB18388.1; JOINED.
 DR EMBL; 066771; AAB18388.1; JOINED.
 DR EMBL; 066772; AAB18388.1; JOINED.
 DR EMBL; 066773; AAB18388.1; JOINED.
 DR EMBL; 066774; AAB18388.1; JOINED.
 DR EMBL; 066775; AAB18388.1; JOINED.
 DR EMBL; 066776; AAB18388.1; JOINED.
 DR EMBL; 066777; AAB18388.1; JOINED.
 DR EMBL; 066778; AAB18388.1; JOINED.
 DR EMBL; 066779; AAB18388.1; JOINED.
 DR EMBL; 066780; AAB18388.1; JOINED.
 DR EMBL; 066781; AAB18388.1; JOINED.
 DR EMBL; 066782; AAB18388.1; JOINED.
 DR EMBL; 066783; AAB18388.1; JOINED.
 DR EMBL; 066784; AAB18388.1; JOINED.
 DR EMBL; 066785; AAB18388.1; JOINED.
 DR EMBL; 066786; AAB18388.1; JOINED.
 DR EMBL; 066787; AAB18388.1; JOINED.
 DR EMBL; 066788; AAB18388.1; JOINED.
 DR EMBL; 066789; AAB18388.1; JOINED.
 DR EMBL; 066790; AAB18388.1; JOINED.
 DR EMBL; 066791; AAB18388.1; JOINED.
 DR EMBL; 066792; AAB18388.1; JOINED.
 DR EMBL; 066793; AAB18388.1; JOINED.
 DR EMBL; 066794; AAB18388.1; JOINED.
 DR EMBL; 066795; AAB18388.1; JOINED.
 DR EMBL; M59832; AAA63215.1; -
 DR PIR; A35893; MMHDMH.
 R HSP; 060675; 1000.
 R GeneW; HGNC:6482; LAMA2.
 DR MIM; 156225; -
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001886; LamNT.
 DR InterPro; IPR000034; Laminin_B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00052; Laminin_B_2.
 DR Pfam; PF00053; Laminin_EGF_15.
 DR Pfam; PF00054; Laminin_G_5.
 DR Pfam; PF00055; Laminin_Nterm_1.
 DR PRINTS; PR00011; EGF-LAMININ.
 DR ProDom; PD002083; LamNT_1.
 DR ProDom; PD003031; Laminin_B_2.
 DR SMART; SM00180; EGF-Lam_14.
 DR SMART; SM00001; EGF-like_3.
 DR SMART; SM00281; Lamb_2.
 DR SMART; SM00282; Lamb_5.
 DR SMART; SM00136; LamNT_1.
 DR PROSITE; PS00022; EGF_1_11.
 DR PROSITE; PS01186; EGF_2_3.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF_14.
 DR PROSITE; PSS0025; IAM_G_DOMAIN_5.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Collid coll;
 Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 3110 LAMININ ALPHA-2 CHAIN.

FT DOMAIN 23 286 LAMININ N-TERMINAL (DOMAIN VI).
 FT 287 343 LAMININ EGF-LIKE 1.
 FT 344 413 LAMININ EGF-LIKE 2.
 FT 414 468 LAMININ EGF-LIKE 3.
 FT 469 517 LAMININ EGF-LIKE 4.
 FT 518 527 LAMININ EGF-LIKE 5 (N-TERMINAL).
 FT 528 723 LAMININ DOMAIN IV 1 (DOMAIN IV B).
 FT 724 756 LAMININ EGF-LIKE 6.
 FT 757 806 LAMININ EGF-LIKE 7.
 FT 807 864 LAMININ EGF-LIKE 8.
 FT 865 917 LAMININ EGF-LIKE 9.
 FT 918 966 LAMININ EGF-LIKE 10.
 FT 967 1013 LAMININ EGF-LIKE 11.
 FT 1014 1059 LAMININ EGF-LIKE 12.
 FT 1060 1105 LAMININ EGF-LIKE 13.
 FT 1106 1165 LAMININ EGF-LIKE 14 (N-TERMINAL).
 FT 1166 1175 LAMININ EGF-LIKE 15 (C-TERMINAL).
 FT 1176 1379 LAMININ EGF-LIKE 16.
 FT 1380 1419 LAMININ EGF-LIKE 17.
 FT 1420 1468 LAMININ EGF-LIKE 18.
 FT 1469 1526 LAMININ EGF-LIKE 19.
 FT 1527 1573 LAMININ EGF-LIKE 20.
 FT 1574 2144 LAMININ II AND I.
 Query Match 6.28; Score 186.5; DB 1; Length 3110;
 Best Local Similarity 19.9%; Pred. No. 1.4;
 Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;
 QY 21 NNOIKKEET-----NNYDIEKRSDDREESTINVENENATFLOKTPODTHLLEE 71
 DB 1693 NEKAIKLNTLGTFRDEAFERNLEGLQK-----EIDOMKELRKKLTOKEIATED 1742
 QY 72 EVKSSSVESNSSIDIAOOPSHTTN---REESVQTSNVEDS-HVSDFAKSIKESN- 126
 DB 1743 ELVAEHLKKVKKLPESGENEMEKDLREKLADYKKNVDDMDLRETDRIENR 1802
 QY 127 -----TESGKE-ENTIEQPNKV-----KEDSTTS 149
 DB 1803 LFAVNQKNMTALEKKKEAVESGKROENTLEKGDIIIDEANRLADEINSIIDYVEDIOTK 1862
 QY 150 QPSYTNIDKESN-----ODELLNLPINEYENKARPLSTSA-----QPSIKRVYVQL 199
 DB 1863 LPPNSELNDKIDLSOEIDRKLAEKVQSNESAQAQNDSSAYLDGLDEAKAINFNAT 1922
 QY 200 AAEQG-SNVNHLIKVTDQSITEGYDSEGVYKADAENLIYDVEYDVKVSGDTMTVD 258
 DB 1923 AAFKAYSNINDYI-----DEAEKVAK--EANDLAHEAT-----KATGPR--G 1961
 QY 259 IDKNTVPSDLTDSFTI-----PKIKNSGEIATGT---YDNKKOITTYFTDYVD 306
 DB 1962 LLKEDAKGCLQGSFRILNEAKKLANDYKEMEDHNLGLTRIENADARGDLRLTNDLGL 2021
 QY 307 KYENI--KAHLKTSYIDSKVVPNNNTK-----LDVEKKTALLSNVKTIT 349
 DB 2022 KLSAIPNDTAAKLDVADKAAQANDTAKDYLQITELHQNDGLKKNYKRLADSVAKTNA 2081
 QY 350 VEYQRPENR-TANLQSMFTNID-----TNHYTEQITTYINPLRYS 389
 DB 2082 V-VKDPKKNKIADADATVKNLEQADRLIDKLPKLEQDNKKNISEIKELINQARKQ 2140
 QY 390 AKETNVNISGND-----EGS-----TIIDSTIIVK-----YK 418
 DB 2141 ANSIRKVSQSGDCIRTRYKPEIKKGSYNNIVNVKTAVDADHILFLYLGSAKFIDFLAEMR 2200
 QY 419 VGDNONLPDSNRIRYDYSIEDVTNDD-----YAOGNNDNVNINPNDISPIYIKYS 471
 DB 2201 KGKVSFLMDVGSVGVRVEYPLDTTIDDSYWRIVASRTGRNCTISVR--ALDGPASATPS 2228
 QY 472 -KYDPNKDDYTTIOQTYWQITNEYTEGFRTASVDNTIASTSGOGGDLPEKTYKI 530
 DB 2259 THHSITSPGGYTIILDVDANAMLFVGLTQKLLKADAVRITFTGCGGEIYFPNKP-----I 2313
 QY 531 GDYVMEVDVKGIGIONTNDNEKPLSNVLTLYTPD--GTSKSVRT--DEGKYQFQDG 582

```

Db 2314 G-LMNREKES-----DCKGCTVSPQVDSGCTIQFDG 2345

RESULT 10
TANX_XENLA
ID TANX_XENLA STANDARD; PRT: 1744 AA.
AC 001550:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanabiti.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=radpole head.
RX MEDLINE=92398961; PubMed=1524825;
RA Hammati-Britvanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate
RT neurons defines a new class of intermediate filament protein."
RL Neuron 9:417-428(1992).
CC -1- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURULA AND PERSIST
CC DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
CC CORD.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M99387; AAA49966.1;
DR PIR; JH0720; JH0720.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament_1.
DR PROSITE; PS00226; IF_1.
DR Intermediate filament; Coiled coil; Neurone.
FT DOMAIN 1 12 HEAD.
FT DOMAIN 13 314 ROD.
FT DOMAIN 315 1744 TAIL.
FT DOMAIN 8 48 COIL 1A.
FT DOMAIN 49 60 LINKER 1.
FT DOMAIN 61 156 COIL 1B.
FT DOMAIN 157 179 LINKER 12.
FT DOMAIN 180 193 COIL 2A.
FT DOMAIN 194 199 LINKER 2.
FT DOMAIN 200 314 COIL 2B.
SQ SEQUENCE 1744 AA; 199561 MW; 6502EAC9F66C4E93 CRC64;

Query Match 6.1%; Score 184.5; DB 1; Length 1744;
Best Local Similarity 19.7%; Pred. No. 0.86;
Matches 149; Conservative 122; Mismatches 239; Indels 245; Gaps 42;

OY 1 SDEKKNVNNNOSINTDNNQIKKETTNN-YDGIEKRSERTESTTNDENAEPL 58
DB 886 SFOIKDFEINQECTKSPQIREAPDTEVDHQVDFMQGSEFEREQGLNNIKQEVDTL 945
OY 59 QKTPD-----NTHLTEREVK-----ESSSV-----ESSNSIDP- 88
DB 946 QWYDEDSFQNNDEPOLESQDQEQKIKLEENQJSENGNONGFNGNDIEFSQOQYDID 1005
OY 89 -----AQQPSHTINRE-----SVQSDNVEDSHVDFANSKIKESN 126
DB 1006 EICQETIGNVSAQLCESDINQDKLSMEDDEEQNNPFTEDNIGLEQSDQENTRNEGT 1065

```

```

OY 127 TESGKEENTIECPNKKEDSTTSQPSGYNTIDEKISNO--DELLNLPINE--YENKARPL 182
DB 1066 KFSQECDCVYKRP-----ED--MSDKSEYSGQOEDLDKQVDTSLNMQANLLEKEEYIL 1119
OY 183 STTSAQPSI-----KRVYNQALAEQGSVNVHLIKVTDQSI--TEGY----- 222
DB 1120 HHADQRSVNDIEITIDEKLSERIIDLNELAT--VDVNSLANKQVDFLDEYAVDDNV 1176
OY 223 ---DSEGVYKAH-----DAENLYDVTFE-----YDKVKSQGTMYVDI 259
DB 1177 GMQDDDSQOYQKEDLFVDGNNIIKIEIQCTSLNLOEICEERVMDVDEDI--SGEAKNESV 1235
OY 260 DKNTPV-----SDLTDS-----FTPIKIKNSGRIANGTYDNKNKQITTYF 301
DB 1236 ENMNDVVDLPEAKVYTGDEQISPLQDEKLNLEMTEDTKDNDGOL--CLEKENE----- 1285
OY 302 TDYVDKYENIK-----AHLKTSYIDSKVY--NNNKL-----DVEYKTAALS-- 343
DB 1286 TEYIEVTSPOCATLSDHAGREL--VDQNSANLOFCENPRTKTLAHHIEVETVAADSL 1343
OY 344 -----VNKTIVEXQRPENETANLQSMFTNIDFKNHTVEQTIYINPLYSAKETNVI 397
DB 1344 ESTEROVETERIRP-KPEDSKMEN-----ENSESESVDSQEISLN--SHKSEEFEEI 1393
OY 398 SGNQDEGSIIDDSITIKYKYGDNQNLPSDKRIYD--SEYEDVYN--DDYAGLNNNDY 454
DB 1394 S-----KDYOL--EQLPDPVPLPLDEFDLTPQPVHEHONDD 1435
OY 455 NINFGNIDSPYIKYISKYDPNKKDYTTIQCTVTQMOTINETY-GEFRTASDNTIAFST 513
DB 1436 GA-----STFTISV-----DEDKEREYRESVSDNESENEEPGDVLSVDTSQVEYIT 1483
OY 514 SSGQGGDLPPEKTYKIGDYVEDVYDKGIQNTND-NEKPLSNVL----- 557
DB 1484 LSGLAQ-----PPSY-LGD--NESESDSMENAEILNENPSDIYDFVNSQMTETKIIIA 1534
OY 558 -----VTLVYPDQTSK--SVRTDEDKXQOFD 581
DB 1535 EQVTEQTEVTLQFDPAFNKLTENLAREKETDYVE 1569

RESULT 11
YGA4_YEAST
ID YGA4_YEAST STANDARD; PRT: 817 AA.
AC P46949;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 90.8 kDa protein in RRP41-SNG1 Intergenic region.
GN YGR196C OR G7589.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96076633; PubMed=7502584;
RA Guerrero I., Mala e Silva A., Barreiros T., Arroyo J.,
RA Garcia-Gonzalez M., Garcia-Saez M.I., Rodrigues-Pousada C.,
RA Nomela C.;
RT "The complete sequence of a 9000 bp fragment of the right arm of
RT Saccharomyces cerevisiae chromosome VII contains four previously
RT unknown open reading frames."
RL Yeast 11:1087-1091(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```


CC -----
 DR EMBL: X82775; CA58019.1; -
 DR EMBL: Z72981; CA97222.1; -
 DR SGD: 50003428; YGR196C.
 KM Hypothetical protein.
 SO SEQUENCE 817 AA; 90797 MW; E52C5D659D63BEBB CRC64;
 Query Match 6.1%; Score 184; DB 1; Length 817;
 Best Local Similarity 21.0%; Pred. No. 0.37;
 Matches 134; Conservative 88; Mismatches 237; Indels 180; Gaps 31;
 Y 1 SDEPKNDVNN-----NOSINFDNNNOI--IKKEFNNDGIEKSED 42
 Db 120 SSEHSKSDYLSTSLKSPSENKSPHTRNRAVNDLNLQISREMP--IKQSD 176
 Y 43 RTESTNVDE--NEATFLQKTPQDNTLHTEEVKSSSVSSNSIDTAQPSHTTINRE 100
 Db 177 RRDSDSCDEIONAPLGEAVPSSSPVEDEKSHSLGVSMDTNEADPTF--NPTFRGN 233
 Y 101 ESQVTSQNVESHVSDPANSKIKESN-----TESGKEENTIEQPNKVKEDSTISQPSGYT 155
 Db 234 EHLSSDGVSSQKODEF--KVSERGLADILPAKEEN-LQO-----EDGGEVSSGAL 284
 Y 156 NIDKISMODELLNLPINEYENKARPLSTSAQPSI--KRVYNQOLAEOGSVNNHLIK 212
 Db 285 EKKEKSEKTSIRN-----RN-----STSGQDKVAKPKRVANETKSDNGYNSFFND 333
 Y 213 VTDSITEGYDSEGVAKAHDAENLIYDVFPEVDKYSQDPTVVDIDKNTVPSDLAD-- 270
 Db 334 YQHSDDSEEDNNNEGSSSD-----DDNRSVSADKADIRKQKOLDTDD 381
 Y 271 --STTIKIKNSGEIITATGYDNKKQIYTFDYDKYENKAKHLKLSYIDKSKVYN 328
 Db 382 ALSTY-ESIKYSTETEEDNEDESIE-----DKNE-----DNESIED 419
 Y 329 NNTLADVEYKATLSVNTIYEVQRPEN-----RTANLQSMFTNIDKN 374
 Db 420 ENEDTD-SYKFSNREKSGILTSDEEEKGMSSDSDEGSKAPSGYFSKMIG-DDKG 477
 Y 375 HTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDSTIIVKYGVGNQNLPSNRIVDY 434
 Db 478 DSALQPNQIDPT-----ENTNLSNGSELENSDDEBDHINECKYLESSVAKDSTVDWSW 533
 Y 435 SEYEDVTMDYAOUGNNDVNIINGNIDSP--YII-----KVISKYD-- 474
 Db 534 KPDSALRSGVQ-----DTANKKAPPGYVIDSNGKLVDLTPASMKRIVYSTSEM 584
 Y 475 -----PNK---DDYTIQOTVPMQTTINETYGEFRASYDNTIAFSTSGOG--QGDLP 523
 Db 585 ESTMDAFPSKEDDLETTIRDTKIT-----YDNTTIYVPGILGNQMLP 629
 Y 524 P-----EKTYRIGDYVWEDVDKGIQNTNDRKPLSNVL 557
 Db 630 PLPMDAQEQNLNAGN-----DNSTDNDSNNTANDL 660
 RESULT 12
 RBP2_PLAVB STANDARD; PRT: 1251 AA.
 ID RBP2_PLAVB
 AC 000799;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Reticulocyte binding protein 2 (Fragment).
 GN RBP2.
 OS Plasmodium vivax (strain Belen).
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CC NCBI_TaxID=31273;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92315383; PubMed=1617731;
 RA Galinski M.R., Medina C.C., Ingravallo P., Barrwell J.W.:
 "A reticulocyte-binding protein complex of Plasmodium vivax

RT merozolites".
 RL Cell 69:1213-1226(1992).
 CC -FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
 CC HUMAN RETICULOCYTE CELLS.
 CC -SUBCELLULAR LOCATION: Membrane-bound (Probable).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M88098; AAA29744.1; -
 KW Malaria; Receptor; Membrane.
 FT NON_TER 1
 FT NON_TER 1251
 SO SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;
 Query Match 6.1%; Score 184; DB 1; Length 1251;
 Best Local Similarity 20.0%; Pred. No. 0.61;
 Matches 140; Conservative 115; Mismatches 236; Indels 210; Gaps 37;
 Y 11 NNNQINPTDNNQIJKKEE--TNVYDGIKRSERTE-----STNV--DENATFLQK 60
 Db 302 SNNKVNINENIRNSEOYLKIDAEKQASTKVELFKHETTISIFKESELIGYET 361
 Y 61 TPQDNTLHEEVKSSSVSSNSIDTAQPSHTTINREESVQTSQNVESHVSD-FAN 119
 Db 362 KSQKIKKAKEDIMKE--IERHNEIQOVAGPOENLNKLNKPNHYCAEDELNNDKSTN 418
 Y 120 SKIK-ESNTESGKEENTIEQPNKVKEDSTISQPSGYTNIDKISQDELLNLPINEYENK 178
 Db 419 AKVLETLNESVK--HNLSEITNIKOG-----EKIYSK-----AKDIMQK 457
 Y 179 ARPSTLSAQSIRKRVYNQOLAEOGSVNNHLIKYTQDS---IIE--GYDOS---E 226
 Db 458 IKATSEMTAEKTEKLV-----KDDQSNVYVNLQTTETERNLIVTEKNLNGIDSTIYIE 512
 Y 227 GVIKAHAENLIYVTF-EVDKYSQDPTVVDIDKNTVP-----SDLTDSFTPIKIKD 279
 Db 513 GALKESKGN--YEIGFLEKLEBEIGKNNKLVYDIKKKINSTGVNFSLPFNFDLQYDF 569
 Y 280 NSGETIATGYDNKKQIYTFDYDKY-ENIK-----AHLKLSYIDSKYV 326
 Db 570 NKN-----INDYENKKGELIYNEFEGSLINKISELIRNASSENTSDYNSAKTLRLAQRKYNVL 625
 Y 327 PN-----NNTKLADVE-----YKATLSVNTIYEVQRPENTANLQSMFTNID 371
 Db 626 LNKEEANKYLRDVAKKVESFRIFENMKESLDKINEMIKKEQLYVNEGH-GNVKQLVENIK 684
 Y 372 T-----KNHTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDSTIYIK 415
 Db 685 ELVDENNLSDILKQATGNNEI-QKITHSTLKNKKK-----TILGHVQDSAKVGG--IK 735
 Y 416 V-----YKVGNOQLPDSNRIVDYSEYEDVTNDYAOUGNNDVNIINGNIDS 463
 Db 736 ITPELATLTELGDADAKLTAQDELKPESKNNVVLEFENMSK-----NTNELDVH-KNIOD 787
 Y 464 PY--IIVKISKYD-----PNKDDYTIQ--QIVTMQTTINETYGEFR-- 502
 Db 788 AYKVALETLAHSDEIDTWKQDSKSLIEGNQIYLVLLNQKNNISSIKSKEAPVSKI 847
 Y 503 -----ASYDNTIAF--STSGQOCGLDPEPKYKIGDYWEDVDKGI 543
 Db 848 GNVSKKHSELKITSQDSYDNIILAEKOTELQNLNRSFTQET----- 891
 Y 544 QNTNDNEK-----PLSNVLTLYPDGTSKSVRTDED 575
 Db 892 -NTNDSKLEIKITDFESLKNALKTL--EGEVNALKASSD 928


```

CC -1- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X16561; CAA34560.1; -.
CC PIR: S07485; RNZQ2L.
CC InterPro: IPR000684; RNA_PolII_repeat.
CC InterPro: IPR000722; RNA_Pol_A.
CC InterPro: IPR002879; RNA_Pol_A2.
CC Pfam: PF00623; RNA_Pol_A; 1.
CC Pfam: PF01854; RNA_Pol_A2; 1.
CC PROSITE: PS00115; RNA_Pol_II_REPEAT; 9.
CC Transferase: DNA-directed RNA polymerase; transcription; zinc; Repeat;
CC KM DNA-binding; Nuclear protein; phosphorylation; zinc-finger.
CC FT Z-FING 68 84
CC FT DNA_BIND 378 411
CC FT DOMAIN 707 725
CC FT DOMAIN 1093 1128
CC FT DOMAIN 1144 1159
CC FT DNA_BIND 1182 1193
CC FT DOMAIN 1687 1694
CC FT DOMAIN 1258 1290
CC FT DOMAIN 1261 1290
CC FT DOMAIN 1602 1612
CC FT DOMAIN 1746 1759
CC FT DOMAIN 1806 1820
CC FT DOMAIN 2061 2246
CC FT DOMAIN 2247 2384
CC SQ SEQUENCE 2452 AA; 278166 MW; F995E117617A48F CRC64;

Query Match 6.1%; Score 183; DB 1; Length 2452;
Best Local Similarity 21.1%; Pred. No. 1.5; Mismatches 275; Indels 172; Gaps 28;
Matches 145; Conservative 95;

3 DEERNDVYNNQSIINTDNNQIIRKEETNNYDGIKESDRTSTSTVNDENAEFLQK-- 60
Db DDDLDDEENISDNINIGNRKYNGTILKNND-----ENMLNPIDVHKVNNFLEKLV 1336
QY 61 -TPDDNTHLEEEVKESSSVSSSIDTAAQPSHTTINREESVQTS-----DN 108
Db 1337 IIRKINSN-----DRLSYEAQNNATILKLAHLRYLNSKLLQTRHKVSKGLDWLQ 1389
QY 109 VE-----DSHVSDFANSKIKESNTES-----GKEENTTECP-----NKV 142
Db 1390 IEKIFYISLCPGRCVGLAAQSIGEPATQTLTFPHFAGVGSKNVLYGVRKLKELINIV 1449
QY 143 KEDSTSPSGSYTINDEKISNDELNLPINEYENKARPLSTSAOPSIKRVYNOQLAAE 202
Db 1450 KNVNT-----PSTTIYLDVNSWDOOKADILKLETTILKQLTSHAQIITDNTTITIEE 1506
QY 203 OGSVNVHLIKVYDOSITEGYDSDSGVIRKADAENLIYDVFEVDKVSQDTMTVDIDKN 262
Db 1507 DKSVMNEPEEPDDDTQ-YSLGEMVLR-----IQLTINHVNEK-KLTAKELIYIYS 1557
QY 263 TVPDDLDLSFTIRIKIKNSGSIIT-----ATGYDNKNKQITTYTDDYDKENIRAKL 315
Db 1558 VFSSDELD--IYTDNSEDLVLRIRVKYLNGEYNFNMVDVYDVAEVODEDEDEHL 1614
QY 316 -----KLTYSYDKSKVPPNN-----NTR-----LDVEYKTAALS 342
Db 1615 VANDRGVDETCKNSTHPIHHDYNNMTTINFKSKVKNRNISSDINTKNEDSISINSSNDEQV 1674

```

```

QY 343 SYNKITVEQRPENENTANLQSMFTIDKKNHVE-----QTIYNPLRYS 389
Db 1675 NISSPVSNNMHHNNNNNNNNNDSSNTINDIKYKNIKKEGNECALRGGDSNTSALFGKNS 1734
QY 390 AKETNV--NISGNDESGTIIDSTIKYKVGDNQMLPSNRIRYDYSEVEDYND-- 444
Db 1735 OKEDNIYNNNNNDND--DEEEEDLFGDHANSPKNTKGKKNKNNKNSNNENK 1790
QY 445 YAOGLNNNDVNNF--GNIDSPYIIRVSKYDPKNDYTTIQTYMTQTTINEYGEFR 501
Db 1791 NKSGNNNSNNSTYDDGDVDN-----DNDNDNDNKSQDI-----TIKE----- 1829
QY 502 TASYDNITAF-STSSGCGGGLDPEKTYKIGDYWEVDYDKDQIONT--NDNEKPLSNVL 557
Db 1830 ----DNDVAEMKSTKNAEIDL-----ELKKNHIEHISREDTDTFLKLMEOCLSTLK 1881
QY 558 -----VTLTYPDGTSKSVRTDEDKY 578
Db 1882 LRGIENITKVMRESKITYDSNGKF 1908

RESULT 15
YCG1_YEAST STANDARD; PRT: 853 AA.
AC P25588; P25589; P27513; P87003;
DT 01-MAY-1992 (Rel. 22, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 97.9 kDa protein in CHA1-KRRI intergenic region.
GN YCLO61C OR YCLO61C/YCLO6C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid:4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Rasmussen S.W., von Wettstein D.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Gromadka R.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 384-596 FROM N.A.
RC STRAIN=ATCC 28383 / FL100;
RX MEDLINE=90384830; PubMed=2169608;
RA Kern L.;
RL "The URA1 gene of Saccharomyces cerevisiae encoding uridine kinase.";
CC Nucleic Acids Res. 18:5279-5279(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59720; CAA42405.1; -.
CC EMBL: X53998; -. NOT_ANNOTATED_CDS.
CC PIR: S19391; S19391.
CC PIR: S19391; S19392.
CC PIR: S21360; S21360.
CC PIR: S29373; S29373.
CC SGD: S0000566; YCLO61C.
CC Hypothetical protein.
CC CONFLICT 505 505 L -> V (IN REF. 3).
CC CONFLICT 567 567 MISSING (IN REF. 3).
CC SEQUENCE 853 AA; 97946 MW; 16E09FCC0BF248D1 CRC64;

Query Match 6.1%; Score 182.5; DB 1; Length 853;
Best Local Similarity 18.9%; Pred. No. 0.45;

```

Matches 121; Conservative 117; Mismatches 234; Indels 169; Gaps 29;

```
QY 5 EKNDVNNNSINTDNNQ---IKKEETNNYDGIKRSERDRETTNVDENEAFLQK 60
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 291 EENFOLNAHSDSGSDSGFALSGNEIADYESGSENNRRESDSEKEDDEILKQK 350
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 TPQDNTL-----TEEEVKESSSVSSNSITAOQPSHTTN---REBSVOTSDNV 109
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 351 KSHVKKHIIINESDSDTEVEAKPKKEADES-----LPKRIAINLGHYGDNIGEDTDKF 402
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 110 FDSHVSOEPAKIKESNTESGKEENTIEQPNKVEDSTSCPSGVTNDEKISNO---DE 166
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 403 QETVLDPTQNT-----EVAERTITL-NEVKDDVYNEDA-----DEAIRQLIDKE 449
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 LLLNPINEYENKARPLSTSAQPSIKRYTVN--QLAEQGSNNHLIKVTDQSITEGYD 224
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 450 KLOLKOKEKEHEAK-----IKELKKRGVTNPFEMEAESEDEMHGIGADGEGSDYDS 503
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 225 -----SEGVIKAHDAENLIYDVFVDKYSKGDVTVDID-----KNIVPS 266
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 504 DLEKIIDYSKNNFNPHEIREMLAENKEMD--IKMINKIIVYDIKNGGFRNRKRAKNSLEL 561
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 267 DLTDSFTIPKIKDSGEIATGTYDNKKQITFTFDYVDKXENIKAKHLKLSYIDKSKV 326
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 562 ELSD-----DDEDDVLQ--QYRLKRRRLMR-----KRRLEIG--DDAKL 596
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 PNNNTKLDVEYKKTALSSVNTKITVEYORP---NENRTANLCSMTNIDTKNHTVEQTIYI 383
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 597 V-KNPKSAFFESMVEDI-----IFKNPFGAEFEYNLDITATDLDT----- 639
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 384 NPLRYSAKETVNIISGN-GDEGSTIID--STIIKYYKVGDNQNLPSDSRIYDYSEYED 439
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 640 -----ODNSINVGDNNTGNNEQKPVQKNNKKVLIISEDVCKSLSLKSN--NYEDFE- 688
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 440 VTNDYQALGNNDVNIENFGNIDSPYIIKVISKYPNKKDDYTTIQOYVIMQTTINEYGE 499
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 689 -TDKELSRIOHGNDCAI-----EDLYTLKONSSIKSFTNSQT-D 725
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 500 FRTASDYDNTIAFSTSSGGGGLDLPPEKTYIKIGDYVWEDVDKDIQNTNDNE----- 550
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 726 STSKTVNTIIDLKRPEDDEDEVENGDTSYGVFKHPSIISKFASTRIDINDKFKEGNKTV 785
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 551 -----KPLSNVLTLYTPDGTSKSV---RTDEDGKYQFD 581
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 786 KILKSYKTVGSSKASITYMGKTRKLIAPKRKTEGSHRYHHD 826
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 14, 2003, 12:59:00
Job time : 19.6487 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 22.2509 Seconds
(without alignments)
5389.420 Million cell updates/sec

Title: us-09-147-405b-13

Perfect score: 3013
Sequence: 1 SSDEKNDVYNNNSINTDD.....PDGTSKVRTEDEGKYQFDG 582

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2	070022 staphylococ
2	2810.5	93.3	931	2	09KI13 staphylococ
3	1160.5	38.5	1171	2	09KW6 staphylococ
4	1144.5	38.0	1166	2	086489 staphylococ
5	1137.5	37.8	1141	16	099W46 staphylococ
6	1137.5	37.8	1141	16	0932F7 staphylococ
7	576	19.1	1315	2	086488 staphylococ
8	559	18.6	1385	16	099W47 staphylococ
9	499	16.6	1733	2	09KI14 staphylococ
10	489	16.2	953	16	099W48 staphylococ
11	460.5	15.3	881	2	093M7 staphylococ
12	455	15.1	947	2	086487 staphylococ
13	452	15.0	933	2	053653 staphylococ
14	452	15.0	935	16	0932C5 staphylococ
15	452	15.0	989	16	099V14 staphylococ
16	397.5	13.2	877	16	099R07 staphylococ

17	396	13.1	913	2	086476 staphylococ
18	376	12.5	940	2	053682 staphylococ
19	344.5	11.4	961	16	099RD3 staphylococ
20	326	10.8	1038	16	099RD2 staphylococ
21	268.5	8.9	1301	5	08W5K5 plasmodium
22	266.5	8.8	1160	2	08RM86 streptococ
23	266	8.8	2402	2	09A8R7 staphylococ
24	261	8.7	1161	2	09X3M7 streptococ
25	259	8.6	1161	2	08RJ10 streptococ
26	239	7.9	1039	2	P72534 streptococ
27	237.5	7.9	2276	2	093YV6 staphylococ
28	234	7.8	604	5	026021 plasmodium
29	231.5	7.7	1401	2	08RJR4 staphylococ
30	228	7.6	2647	5	0904X0 plasmodium
31	225	7.5	1463	2	086919 staphylococ
32	220	7.3	2843	5	0963L5 dictyostell
33	219.5	7.3	1364	5	08T223 dictyostell
34	217.5	7.2	559	5	0903Y8 plasmodium
35	216.5	7.2	1959	16	08YAN1 listeria mo
36	216.5	7.2	3394	5	077384 plasmodium
37	215.5	7.2	1564	2	093IM3 plasmodium
38	215	7.1	518	12	09DH49 amasacta moo
39	214.5	7.1	2081	10	09LH98 mycoplasma
40	212.5	7.1	1125	16	098PM9 plasmodium
41	212.5	7.1	3130	5	098K46 plasmodium
42	212	7.0	2771	5	026216 plasmodium
43	211.5	7.0	1344	2	049545 mycoplasma
44	211.5	7.0	2206	5	096205 plasmodium
45	211	7.0	691	5	09GSD1 plasmodium

ALIGNMENTS

RESULT 1

ID 070022

PRELIMINARY:

PRT: 1092 AA.

AC 070022;
DT 01-AUG-1998 (TREMREL, 07, Created)
DT 01-AUG-1998 (TREMREL, 07, Last sequence update)
DT 01-JUN-2002 (TREMREL, 21, Last annotation update)
DE Fibritogen-binding protein precursor.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB;
RA MEDLINE=98261511; Pubmed=9596732;
RX Nilsson M., Frykberg L., Flock J.I., Pel L., Lindberg M., Guss B.;
RT "A Fibritogen-binding protein of Staphylococcus epidermidis."
RL Infect. Immun. 66:2666-2673(1998).
DR EMBL: Y17116; CA76638.1; -;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor: 1.
DR TIGRFAMS: TIGR01167; LPXTG_anchor: 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal: 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 51
FT CHAIN 52 1092 POTENTIAL FIBRINOGEN-BINDING PROTEIN.
SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AD8B984 CRC64;
Query Match 100.0%; Score 3013; DB 2; Length 1092;
Best Local Similarity 100.0%; Pred. No. 1,9e-105;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSDEKNDVYNNNSINTDDNNQIKKETNNYDIEKRSDEKSTNVDEKATFLQK 60
DB 75 SSDEKNDVYNNNSINTDDNNQIKKETNNYDIEKRSDEKSTNVDEKATFLQK 134
QY 61 TPQDTHLTREEVKKSSVESNSSIDTAQPSHTTINRESVQISDVAVEDSHVSDPFANS 120

```

|||||
Db 135 TPQDNTHLTFEEVKESSSVSSNSSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANS 194
OY 121 KIESNTESGKEENTIEQPNKVKEDSTSPSGYTINDEKISNOBELNLPINEYENKAR 180
Db 195 KIESNTESGKEENTIEQPNKVKEDSTSPSGYTINDEKISNOBELNLPINEYENKAR 254
OY 181 PLSTSAOPSIKRVYVNOQLAEOGSNVNHLIKYTDOSITFEGYDSDGCVIKAHDAENLIYD 240
Db 255 PLSTSAOPSIKRVYVNOQLAEOGSNVNHLIKYTDOSITFEGYDSDGCVIKAHDAENLIYD 314
OY 241 VFEVVDKVKSGDPTMVDIDKNTVPSDLTDSFTIPKIDNSGEIITATGYDNKNKOITYT 300
Db 315 VFEVVDKVKSGDPTMVDIDKNTVPSDLTDSFTIPKIDNSGEIITATGYDNKNKOITYT 374
OY 301 FTDYVVKYENIKAHKLTSTYIDSKVPPNNNTKLDVEYKTLASSVNTTIVEYORPENRT 360
Db 375 FTDYVVKYENIKAHKLTSTYIDSKVPPNNNTKLDVEYKTLASSVNTTIVEYORPENRT 434
OY 361 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGDGSSTIIDSTIIKYYKVG 420
Db 435 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGDGSSTIIDSTIIKYYKVG 494
OY 421 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIITKYSKYDPNKDXY 480
Db 495 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIITKYSKYDPNKDXY 554
OY 481 TTIOQVTVMTQTTINEXYGEFRTASYDNTIAFSTSSGOGGDLDPPEKTYKIGDYVWEDVK 540
Db 555 TTIOQVTVMTQTTINEXYGEFRTASYDNTIAFSTSSGOGGDLDPPEKTYKIGDYVWEDVK 614
OY 541 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 582
Db 615 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 656

```

RESULT 2

```

O9K113 PRELIMINARY: PRT: 931 AA.
ID 09K113
AC 09K113:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibrinogen-binding protein sdrg.
GN SDRG.
S Staphylococcus epidermidis.
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eldhin D., Ilna G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245042; AAF72510.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRPFAMs: TIGR01167; LPTXG_anchor; 1.
DR TIGRPFAMs: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1
SQ SEQUENCE 931 AA: 102955 MW: 591E657D97027116 CRC64:

```

Query Match

Best Local Similarity 93.3%; Score 2810.5; DB 2; Length 931;
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

```

OY 1 SDEEKNDVNNNOSITDNNQIIRKEETNNNDGIEKRSDETRSTTNDENATFLQK 60
Db 74 SSNEEKNDVNNNOSITDNNQ-IRKEETNSNDAIENRSKIDITOSTTNDENATFLQK 132

```

```

OY 61 TPQDNTHLTFEEVKESSSVSSNSSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANS 120
Db 133 TPQDNTHLTFEEVKESSSVSSNSSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANS 192
OY 121 KIESNTESGKEENTIEQPNKVKEDSTSPSGYTINDEKISNOBELNLPINEYENKAR 180
Db 193 KIESNTESGKEENTIEQPNKVKEDSTSPSGYTINDEKISNOBELNLPINEYENKAR 252
OY 181 PLSTSAOPSIKRVYVNOQLAEOGSNVNHLIKYTDOSITFEGYDSDGCVIKAHDAENLIYD 240
Db 253 PLSTSAOPSIKRVYVNOQLAEOGSNVNHLIKYTDOSITFEGYDSDGCVIKAHDAENLIYD 312
OY 241 VFEVVDKVKSGDPTMVDIDKNTVPSDLTDSFTIPKIDNSGEIITATGYDNKNKOITYT 300
Db 313 VFEVVDKVKSGDPTMVDIDKNTVPSDLTDSFTIPKIDNSGEIITATGYDNKNKOITYT 372
OY 301 FTDYVVKYENIKAHKLTSTYIDSKVPPNNNTKLDVEYKTLASSVNTTIVEYORPENRT 360
Db 373 FTDYVVKYENIKAHKLTSTYIDSKVPPNNNTKLDVEYKTLASSVNTTIVEYORPENRT 432
OY 361 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGDGSSTIIDSTIIKYYKVG 420
Db 433 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGDGSSTIIDSTIIKYYKVG 492
OY 421 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIITKYSKYDPNKDXY 480
Db 493 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIITKYSKYDPNKDXY 552
OY 481 TTIOQVTVMTQTTINEXYGEFRTASYDNTIAFSTSSGOGGDLDPPEKTYKIGDYVWEDVK 540
Db 553 TTIOQVTVMTQTTINEXYGEFRTASYDNTIAFSTSSGOGGDLDPPEKTYKIGDYVWEDVK 612
OY 541 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 582
Db 613 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 654

```

RESULT 3

```

O9KMX6 PRELIMINARY: PRT: 1171 AA.
ID 09KMX6
AC 09KMX6:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Bone sialoprotein-binding protein.
GN BDP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=024;
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a
RT member of the staphylococcal sdr family ]."
RL Biochem. J. 345:611-619(2000).
DR EMBL: Y18653; CAB75732.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRPFAMs: TIGR01167; LPTXG_anchor; 1.
DR TIGRPFAMs: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1
SQ SEQUENCE 1171 AA: 127123 MW: C5BC812F9DA5A884 CRC64:

```

Query Match

Best Local Similarity 38.5%; Score 1160.5; DB 2; Length 1171;
Matches 271; Conservative 99; Mismatches 182; Indels 79; Gaps 20;

```

OY 1 SDEEKNDVNNNOSITDNNQIIRKEETNNNDGIEKRSDETRSTTNDENATFLQK 51
Db 1171 SSNEEKNDVNNNOSITDNNQ-IRKEETNSNDAIENRSKIDITOSTTNDENATFLQK 132

```

```

Db 57 STENAKODEASADNKEVSETEENNSTOKNDLTPIKKETNTDSHOEAKKAPTSS7000 116
QY 52 ENENFTLOKTPDNDTHLLEEVES---SSVSESSNSIDTAQOPSHTTINEESVQTSN 108
Db 117 ONNATSTETKPPON--TEKENVKPSTDKTATEDTSVLEEKKAPDNT--NNDVTTKPS-- 170
QY 109 VEDSHVDFANSKIKESNT---ESGKEPNTIEOPNKVED---STTSQPSQTYNI--DEK 160
Db 171 -----TSEIOTPTTPOESTNIENSOPPTPSKVNDQVDTDTTPKEPPVNSKEEL 220
QY 161 TSNDELINLPINE--YENKARPSTTSAPSIRV-----TYNOLAEOGSVNHLLIK 212
Db 221 KNNPEKLELVANDSNTDRSKPPATAPTAPSVAPKRVAKIRFAVAQPAASVANNVNDLIT 280
QY 213 VTQDSITEGYDSEGVIRAHDAENLIYDVTFEVDKVKSGDTMTVVDIDKNTVPSDLTDSF 272
Db 281 VTQKIMEGICID--DGVIOAHGHEHIIYTSDEKIDNAVAGDPTMKVYKXETIPSDITDDE 339
QY 273 TIRKIKNSGEIITAGTYDNKNKOITFTPTDYVDKYEENIKAKHLTSTYIDSKVPMNNTK 332
Db 340 TPVDITDPSGEVIAKGFDLTKFTTKFTDYVDREYEVNNAKLELNSYIDKEVP--NETN 398
QY 333 LDVEYKTALESVNTITVEYORPENRPTANLOSMTNIDTKNHTVEQTIYINPLRSYAKE 392
Db 399 LNLFFATADKETSNAKVEYOKPIYKDESNISQISFSLDTKHEVEQTIYINPLKNAKN 458
QY 393 TNVNI-----SGND---EGSTIIDSTIIKVKVGNQNLPSNRIDYSEYEDVND 443
Db 459 TTVTKSGVANDGYTGDSTIIDSTIISNTEIKYKVASGQOLPOSNRKYDSEYEDVNS 518
QY 444 DYADLGNN---NDVNINGNIDSPYIIKVIKYPNKKDYYTIOQTYVMOTINYEYEEF 500
Db 519 --VTINKYGTNMANNINEGDIDSAIYKVSRYPGABDLAVOOGVMRT---TNKY 571
QY 501 RTASYDN-----TAFSTSSGOGGDL--PEPEYKIGIYVMDVDKDIQNTNDNEK 551
Db 572 NYSSYAGTITLFYQLLTLVYTVSVK-----PEEKLYKIGIYVMDVDKVGQGTDSKX 626
QY 552 PLSNVLTLYTPDGTSKSVRTDEDKYOPDG 582
Db 627 PMANVLTLYTPDGTTSKSVRTDANGHYEFG 657

RESULT 4
ID 086489 PRELIMINARY; PRT: 1166 AA.
AC 086489;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Sdr E protein.
N SDR E.
JS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_Taxid=1280;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., NI Elidh D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RA "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL, AJ005647; CA06652.1;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfams; TIGR01167; LpYTG_anchor; 1.
DR TIGRfams; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SO SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;

```

```

Query Match 38.0%; Score 1144.5; DB 2; Length 1166;
Best Local Similarity 42.6%; Pred. No. 1.4e-35;
Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;

QY 1 SSDEKNDVNNINNOISITDDNNQI--IKKETNNYDGEIKRSEDRTESTNVDENEATFL 58
Db 66 ATTDNKEVSETEENNSTOKNDLTPIKKETNTDSHOEAKKAPTSS7000NNTAT 123
QY 59 OKTPDNDTHLLEEVES---SSVSESSNSIDTAQOPSHTTINEESVQTSNVEDSHVS 115
Db 124 TETPON--TEKENVKPSTDKTATEDTSVLEEKKAPDNT--NNDVTTKPS-----TS 172
QY 116 DFANSKIKESNT---ESGKEPNTIEOPNKVEDSTTSQPSQTYNIDEKIS--NODELNL 171
Db 173 EPSTSEIOTKPTTPOESTNIENSOPPTPSKVND--NOYDATPKPEPVNSKEELNPN 229
QY 172 -----INEYENKARPSTTSAPSIRV-----TYNOLAEOGSVNHLLIKYTDQ 216
Db 230 EKLKELVRANDSNTDRSKPPATAPTAPSVAPKRVAKMRAVAQPAASVANNVNDLITK 289
QY 217 SITEGYDSEGVIRAHDAENLIYDVTFEVDKVKSGDTMTVVDIDKNTVPSDLTDSF 276
Db 290 TIKVG--DGDVYAAAHGDKDIEYDEFTIDNRKVKKGDTMTYNDKNVYPSDLTDKNDPID 348
QY 277 IKDNSGEIITAGTYDNKNKOITFTPTDYVDKYEENIKAKHLTSTYIDSKVPMNNTK 336
Db 349 ITDPSGEVIAKGFDLTKFTTKFTDYVDREYEVNNAKLELNSYIDKEVP--NETSLT 407
QY 337 YKTALESVNTITVEYORPENRPTANLOSMTNIDTKNHTVEQTIYINPLRSYAKETNVN 396
Db 408 FATAGKESQAVTYVDQDPMVHGSNSISITFKLDEKQTIIEQDLYVPLKKSNTFRVD 467
QY 397 ISGNG-----DEGSTIIDSTIIKVKVGNQNLPSNRIDYSEYEDVND--DYA 446
Db 468 IAGSQVDVYDGNIKLNGSTFIIDONTIEIKYKVSNOQOLPOSNRKYDSEYEDVNSO 527
QY 447 OLGNNNDVNINGNIDSPYIIKVIKYPNKKDYYTIOQTYVMOTINYEYGEFTASYD 506
Db 528 KFSNNVATLDFGDINSAYIIKVSRYPGABDLAVOOGVMRT--DKY-GIYVAGYS 585
QY 507 NTIAFSTSSGOGGDL--PEPEYKIGIYVMDVDKDIQNTNDNEKPLSNVLTLYTPDG 565
Db 586 NPIVTSNPTGGGDTGVKREELKYKGDYVMDVDKDGVOGTDSEKPPANVLTLYTPDG 645
QY 566 TSKSVRTDEDKYOPDG 582
Db 646 TTKSVRTDANGHYEFG 662

RESULT 5
ID 099W46 PRELIMINARY; PRT: 1141 AA.
AC 099W46;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
N protein.
OS SDR E SA0521.
JS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_Taxid=158879;
[1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21311952; PubMed=11418146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsunori H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Oi Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka A., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;

```



```

RESULT 7
086488 PRELIMINARY; PRT: 1315 AA.
ID 086488
AC 086488;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE SdrD protein.
GN SDRD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eldhin D., O'Connell D., Cox J., Hook M.,
  Foster T.O.;
RT "Three new members of the serine-aspartate repeat protein multigene
  family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL: AJ005646; CA06651.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK_signal; 1.
SQ SEQUENCE 1315 AA; 14275 MW; 58DAE1F48EE6A689 CRC64;

Query Match 19.1%; Score 576; DB 2; Length 1315;
Best Local Similarity 30.1%; Pred. No. 2,8e-14;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

OY 1 SDEEKNV---INNOSTIDDDNNQIIKEETNNNDGIEKREDETESTNVDEN-NEAT 56
DB 55 STNKELENEATTSASDNQSSDKVDMQOLNODNTRKNOKEMVSSQGNETSNGNKLSIEKE 114
OY 57 FLOKTPQDNTHLTEEEKSSSVESNSSIDTQAQPSHTTINREESVQTSNVEDSHVD 116
DB 115 SVQSTGNKVEVSTAKSDQASPKSTNEDLNTRQ-----TISQEGLO-PDLGKNSYVN 168
OY 117 FANSKIKESNTSGKEENTIEOPNRYKEDSTTSQPSGYNTI-DEKISNODELL--NLPIN 173
DB 169 -----VQPTNEBKRVKVD-----AKTEST-----LNKSDAISKNDDELTVNDNSNS 209
OY 174 EYENKARPLSTTSAPQ-----SIKRVVNOLAEOGSNVNHLIVTQSTIEGVDSDEGYI 229
DB 210 NNEENADITLPKSTAPKRLTRKRIAAVOPSSTEAKNVNDLITSNTTLYVVDADKNNKIY 269
OY 230 KAHDAENLIYDVFEEVDKVRKSGDTMTVIDDKNTV-----PSDLTDSFTIPRIKO-NSG 282
DB 270 PQDYLSLKSQIT--VDKVRKSGDYFTIKY-SDTVQVYGLNPEDIKN--IGDIKPNNG 323
OY 283 EIIATGTYNKKNQITTYFTDYDKYENIKAHKLTSYIDSKSVPPNNNTKLDVEYKTAIS 342
DB 324 ETIATKAKHDTANLLITTYFTDYVDKRNVSQMGINSIYMDADRIIP--VSKNDEFFNVTIG 381
OY 343 SVAKTITVEYQREN--ENRTANLOSMEFTNIDTKNHT-----VEQTIYINPLRYSAKE 392
DB 382 NITTKTTANIQIPDYVYVNEKNSIGSAFT--ETVSHVGNKKNPGYIKOTIYVSENSLTN 439
OY 393 TNVNI-----SGNDEGSTIIDSTIIKRYKVGNDQNLPSNRIYDVS--EYEDVTNDYVA 446
DB 440 AKLQVQAHSSYPNNIGQIKNDVTDIKIYQVPRGYTL--NKGIVDVTKELTDTN-QYL 495
OY 447 Q---LGNNDVINFGNIDSPYIIKYSKDPKDKDYTTIQQVTVMQTTINETGERTA 503
DB 496 QKITTYGDNNSAVIDFGNAASAVYVMVNTKFOYTNSESPTLVQMATLSTSTN-----KSV 549
OY 504 SYDNTIAFSTSSGOGGDLPEPEYKIGDYVMEVDVDKGIQNTNNDKNEPLSNVLTLYTP 563
DB 550 STGNALGFTNNGSGAG-----QEVYKIGNVYMEDTKNKGVDL--GEKYGNTVTVY-FD 602

```

```

OY 564 DGTSKSVR---TDEDGKY 578
DB 603 NNTNTRKVGAVTKEDGKY 620

RESULT 8
099M47 PRELIMINARY; PRT: 1385 AA.
ID 099M47
AC 099M47;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
  protein.
GN SDRD OR SAV0562 OR SA0520.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
  Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohba T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
  Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
  Kanemori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
  Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
  Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
  Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
  Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
  aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003359; BAB56724.1;
DR EMBL: AP003131; BAB41751.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 1385 AA; 149643 MW; 2A0C2D77733B3C1D CRC64;

Query Match 18.6%; Score 559; DB 16; Length 1385;
Best Local Similarity 28.8%; Pred. No. 1,3e-13;
Matches 177; Conservative 117; Mismatches 235; Indels 86; Gaps 27;

OY 1 SDEEKNV---INNOSTIDDDNNQIIKEETNNNDGIEKREDETESTNVDEN-NEAT 56
DB 55 STNKELENEATTSASDNQSSDKVDMQOLNODNTRKNOKEMVSSQGNETSNGNKLSIEKE 114
OY 57 FLOKTPQDNTHLTEEEKSSSVESNSSIDTQAQPSHTTINREESVQTSNVEDSHVD 116
DB 115 SVQSTGNKVEVSTAKSDQASPKSTNEDLNTRQ-----TISQEGLO-PDLLENSYVN 168
OY 117 FANSKIKESNTSGKEENTIEOPNRYKEDSTTSQPSGYNTIDEKISNODELLNLPINEX 175
DB 169 -----VQPTNEBKRVKDAKTESTTTLNVKSDAISK-----NAETLVNDNSNSNN----- 211
OY 176 ENKARPLSTTSAPSIKRVVNOLAEOGS-----NVNHLIKYDQSTIEYDSEGYKA 231
DB 212 ENNADITLPKSTAPKSLNTRMAAIAOPNSTDSKNVNDLITSNTTLYVVDADNSKTIYVA 271
OY 232 HDAENLIYDVFEEVDKVRKSGDTMTVIDDKNTV-----PSDLTDSFTIPRIKO-NSGEL 284
DB 272 QDYLSLKSQIT--VDKVRKSGDYFTIKY-SDTVQVYGLNPEDIKN--IGDIKPNNGER 325
OY 285 IATGYDNKKNQITTYFTDYDKYENIKAHKLTSYIDSKSVPPNNNTKLDVEYKTAISV 344
DB 326 IATAKHDTANLLITTYFTDYVDKRNVSQMGINSIYMDADRIIPD--KKVDPSPSVIGNO 383
OY 345 NKITTYEYQRP--NENRTANLOSMEFTNIDTKNHT-----VEQTIYINPLRYSAKETN 394

```

```

Db 384 ITTTADITTPAYKREADNNSIGSAFT--ETVSHVGNVEDPGYVNOVYVNPMDKDLKGA 441
QY 395 VNISGNDEGST----IIDSTIIIVYKVGDNQNLPSDNRIYDY--SEYEDVTDN--DYA 446
Db 442 LKVEYHPKYPPTINQIOMNTNIRKIRVEGYTL---NKGYDVNTNDLVDVTEDEFKNKM 498
QY 447 QUGNNNDVNINGINDSPILIKIVISKYDKNDDYTTIOQTVMQTTINEYGEFTASID 506
Db 499 TYGSNOSVNLDFGDTISAVVNVNTRKFOYTENSEPTLVOMATELSTGN-----KSVTG 552
QY 507 NTIARSTSSGOGGDLPEPKTYKIGDYVWEDVDKGIQNTNNEKPLSNVLTLYTPDGT 566
Db 553 NALGFTNNQSGAG---QEVYKIGNYWMEDINKNGVQEL--GKGVGNVTVTV-FDNNT 605
QY 567 SKSVR---TDDEGKY 578
Db 606 NTKVGEAVTKEDGSY 620

RESULT 9
Q9K114 PRELIMINARY; PRT: 1733 AA.
ID Q9K114;
AC Q9K114;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative cell-surface adhesin SdrF.
GN SDRF.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CX Staphylococcus.
NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9491;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.U., Hook M.,
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245041; AAF72509.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
Q SEQUENCE 1733 AA; 184720 MW; D8D62EA1692FD4E3 CRC64;

Query Match 16.6%; Score 499; DB 2; Length 1733;
Best Local Similarity 27.1%; Pred. No. 2.8e-11;
Matches 168; Conservative 104; Mismatches 275; Indels 72; Gaps 25;

QY 1 SSDEKNDVYNNQSIINTDNNQIIRKEETNNYDIEKRSEDRSTSTVWDENEA-TFLQ 59
Db 148 TSTTQODSTREKNNPSLKLNDLSSSTSKTDEHSTKQAGST--NKSVLDTNDSPQSE 206
QY 60 KT-PQDNLHLEEEYKESSEVSNS-----SIDPAQPSHTTTN-REESVQTSN 108
Db 207 KTSQANDNDSTNQAPSQKQDSKPEQKYTKFNDDEPTQDVEHTTTKLKPPSVSTDS 266
QY 109 VEDSHVSDPFAANSKIKESNTESGKEENTIEQPKVKEDSTTSQSGVTNDEKI----- 161
Db 267 VNDK--QDYTRSAV-----ASLGVDSENETEATINAVRDNLKAFSREQINEAITALEALK 320
QY 162 --SNODELNLDP--INEXENKARPLSTSAQPSIKRVYNQALAE--QGSNNHLIKYVD 215
Db 321 DSNNDYGVDTPLALNRSQSKNSP--HKSASP---RMNLSLAAPENSGKANNNDKVKITN 375
QY 216 Q--SITEGVDSEGYIKAHDAENLYDYTFEVDKVKSGDTVTVDIDKNTVPSDLTDSFT 273
Db 376 PTLSTNKSNNHANNVIMPTSNQFNLKANYELDSDIKEGDPTFIKYGQYIRPGGELPAI 435
QY 274 IPIKDNSEIATGTGDKKKKQITTYTFPDYDKYENIKAKHLKLSYIDKSAVPPNNNTKL 333

```

```

Db 436 KTLQBSKDSIYANGVYDKTTNTTTFYTFNYVDYQNTIGSFDLIATPRKRETAINDQNY 495
QY 334 DVEYKTAASSVNTKTTVEYQRPENRNTANLQSMFTNIDTKHTVBOYIYNLRSAKET 393
Db 496 PKEVTIANEYVKKDFIYDGNKKQNTT--AAVANVDVNNKHHDEYVYLLQNNQNPXA 552
QY 394 NVNISGNDEGSTIIIDSTI--IKVYKVDNQLPDS-NRIYDYSEYEDVTDNDYQAL- 448
Db 553 KY-----FSTVANGFIPGEVRYEYTDINAVDSNPDLNNSNKVDYSQAPKVS 604
QY 449 GNNNDVNINFGNI---DSFYIKIVISKYDKNDDYTTIOQTVMQTTINEYGEFTASY 505
Db 605 ADGTAVDINFARSMANGKKYITQAVRPRTGNTVTT--EYMLTRGCTTN--TNDYRGK 660
QY 506 DNTIATSTSSGOGGDLPEPKTYKIGDYVWEDVDKGIQNTNNEKPLSNVLTLYTPDG 565
Db 661 STTVYLYNGSSYTAQGDNP---TYSIGDYWYLDKNNNGVQ--DDDEKGLAGVYTL--KDS 713
QY 566 TSKS---VRTDEGKYQFD 581
Db 714 NNRELORVYTTDOSGHYQFD 732

RESULT 10
Q99W48 PRELIMINARY; PRT: 953 AA.
ID Q99W48;
AC Q99W48;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, Dnae statorprotein-binding
DE protein.
GN SDR OR SAV0561 OR SA0519.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CX Staphylococcus.
NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Imai Y., Hori T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami A.,
RA Mizutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003359; BAB56723.1;
DR EMBL: AP003131; BAB41750.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsirK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
Q SEQUENCE 953 AA; 103292 MW; 729A7169A074A1A5 CRC64;

Query Match 16.2%; Score 489; DB 16; Length 953;
Best Local Similarity 28.2%; Pred. No. 3.7e-11;
Matches 162; Conservative 82; Mismatches 236; Indels 94; Gaps 22;

QY 35 GIEKRSDETESTTVDNDEATFLQKTPQDNLHLEEEYKESSEVSNSISDTQAQPSH 94
Db 45 GHEAKAAEHTNDELQSKNETT---APSENK--TTEKV--DSRLKDKTQTATADQPPV 96
QY 95 TTINREESVQTSNDVSDNSVDFANSKIKESNTESGKEENTIEQPKVKEDSTTSQSGV 154
Db 97 T-----MSDSATVKEETSSMQS-----PONATASQSTTQTSNV 129

```

OY	155	TNIDEK----	ISNODELINLPINYEEN-KARPLSTSAOSIKRYVYNOLAA-EQGSNVNH	209
Db	130	TTNDKSSTYSSTNSDEKSNL--TQAKVVSTTPPKTKTHIKORALNMVAINTVAABOQGNTNVD		187
OY	210	LIVATDOST-----	-TEGVDSDSGVAKHADAENLTVDVFEEVDKVKSGDFTW	256
Db	188	KVHFETNLIDLAIDKGHNKTKTGNETFMATSSDYLK-----	LKANVTITDDSVKEGDFTT	239
OY	257	VVIDKNVTVPSDLTDSEFTPIKDNSEGEIATGTGYDNKNKOIYTFTFDYDKYENIKAHLK	316	
Db	240	FHYGOYRPSVSRLPQSOTQLVLYAAGCIIIAKLGIYDSKTNTTYTFPNINYDOQTWNVSGSPE	299	
OY	317	LTSYIDSKVYPNNNTKIDVEYKTAALSSVNTTIVEYQRPRENFETANLOSMTNIDTKNHT	376	
Db	300	QVAFAPKRENAATTOKTAKMEVTELGNDPTYSKDVIADY---GNCKGOOLISTSTVINNEEDS	356	
OY	377	VEQCIYYIN-PLRKSAKETNY-NISGNGDEBSTIIDOSTLIKKYKVAGDONCLPPSNRIYLY	434	
Db	357	RNMVTVYNQKRKYTKETEFTVTLT----GKFNPDKNFKILEVTDQONFVDS-FTPDT	410	
OY	435	SEVEDYTN-DYAQLGNNDVNINFGN---IDSPYIIKVISKDPNKDDYTTIQOTVM	489	
Db	411	SKLDKDVIGQDVIYSNKNRAATYDLNLGSSSQKOYIIIOQVAPFDMSSPDNGKIDVLEET	470	
OY	490	QTTINEYTGFEFRASDYNTAIATFSSTGGOGGLDLPPEKTYIKIGCYWEDVDKDISQIONTON	549	
Db	471	ONGKSSWSN-----SYSNVNGSSSTAND-----	OKKYNLGGYWEEDTKDKGODA--N	516
OY	550	EKPILSNVLVLTLYPDGRS-KSVRFDEDCGXQPOFG	582	
Db	517	EKGIKGVYVILKDSNGKELDRITTTDENGKIQFTG	550	
 RESULT 11 093MH7 PRELIMINARY; PRT; 881 AA.				
AC	093MH7;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Receptor Fbl precursor.			
CN	FBL.			
OC	Staphylococcus lugdunensis.			
CC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;			
OX	NCBI_TaxID=28035;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=2342;			
A	Nilsson M., Ahlen J., Frykberg L., Guss B.;			
.L	"A fibrinogen-binding protein of Staphylococcus lugdunensis.";			
.L	Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.			
EMBL	AF0404823; AAK95649.1; -			
DR	InterPro: IPR001899; Gram_pos_anchor.			
DR	Pfam: PF00746; Gram_pos_anchor; 1.			
DR	TIGRFAMS: TIGR01167; LPRX_anchor; 1.			
DR	TIGRFAMS: TIGR01168; YSIK_signal; 1.			
DR	PROSITE: PS00343; GRAM_POS_ANCHORTING; UNKNOWN_1.			
KM	Signal.			
FT	SIGNAL	1	39	POTENTIAL.
FT	CHAIN	40	881	RECEPTIN FBL.
SO	SEQUENCE	881 AA;	94251 MW;	D4296CA959CAFI9B CRC64;
 Query Match 15.3%; Score 460.5; DB 2; Length 881; Best Local Similarity 27.3%; Pred. No. 3.9e-10; Matches 154; Conservative 88; Mismatches 231; Indels 91; Gaps				
OY	6	KNDVYINNOSITDNDNNOIIEKEEYNNVDTGIEKREDRTTESTTYVDNEATFLQK--TPQ	63	
Db	48	KQOIIGHNDA--TGDT-----QDNKNYN-----	NISSNDEATQQKQIOTS	86
OY	64	DNTHLTEEVEKSSSEVSSNSSIDAQOPSHTTINREESVOGTSDNVEDSHVDFANSKIK	122	

Db	DNVSEAGATNEID-----SHRTYK--ATFALDNNSTLSTOVSPATK	130
QY	87 DNVSEAGATNEID-----SHRTYK--ATFALDNNSTLSTOVSPATK	130
QY	124 ESNTSEK--EEN--TTEOPNKVEDSTSQSGYTNIDEKISN--DELLNPINEYE	176
Db	131 QDITTSNOTTOENNDAFTQTQRTKTKYKODGNMNVLSQVATNDNOSNOPRNSHLNSTVATYN	190
QY	177 N-KARPLSTTSAPQSIRKRVYNOLAEQGSNVNHLIKVTDQSTTEGIDSEGVYKAHDA	234
Db	191 NNHOFRLAKVEANTDNNNTQTSDISKLSNVATIEAAD-----TTPPKA	238
QY	235 ENLIYDTFEVDVYKSGDPTMTVDIDKMTVPDSDLTDSFTPIKIDNSGEIATGTGYNKN	294
Db	239 EYVNLNREFQAPDDVQADSISIKITIPQALNNGVATATAKAPNIMAGD-QIATGTIDEG	297
QY	295 KQITFTFDYDKENIKAHKLTSYIDKSVPNNNTLDEYKTALESVKNKTIVEXOR	354
Db	298 NLI-YTFPDYTNKNNNTIGQISIPGYIDKRVNTHG-KVNLETSIGQTTAKRTVYDEK	355
QY	355 PNEHPTANLQSMFTNIDTKHFTVETQIYINPLRSAKTNNV--ISGN--GDGCTIID	409
Db	356 YGEFRNLISIKGTIDQIDKVNNTYKQTVYVNP-----SSDPTVDPYPLRGSGISDGTNSMVIID	411
QY	410 D-STIKYKKGQNDONLPDSNRIDYSEYEDVNDYDQNLGNNDVHINFG-----NIDSP	464
Db	412 EONTSIKYKVEKAHLDSYV--DPSNYEEDVTSVKITPEKGIYQINFETDDQINSP	470
QY	465 YIKIYSKYDNDKDDYTTIQQVTVMQNTINEYTGCFRTAS--YDNTAFSTSGOGGQDL	522
Db	471 YVYVINGVHDNRNSG-----NLYIRSLTLYGDSFTVSNAMDNVEYHNGSGGDSID	524
QY	523 PPEKTYKIGD-----YVWEDYDKD	541
Db	525 KVPIDPQGDSDSDAYSDSDADSD	548
RESULT 12		
ID	086487 PRELIMINARY; PRT: 947 AA.	
AC	086487	
DT	01-NOV-1998 (TREMBLrel. 08. Created)	
DT	01-NOV-1998 (TREMBLrel. 08. Last sequence update)	
DT	01-JUN-2002 (TREMBLrel. 21. Last annotation update)	
DE	SdrC protein.	
GN	SDRC.	
OS	Staphylococcus aureus.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;	
OC	Staphylococcus.	
NC	NCBI_TaxID=1280;	
RN	SEQUENCE FROM N.A.	
RP	STRAIN-NEWMAN:	
RC	MEDLINE=99098700; PubMed=9884231;	
RX	Joelsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,	
RA	Forster T.J.;	
RT	"Three new members of the serine-aspartate repeat protein multigene	
RT	family of Staphylococcus aureus".	
RL	Microbiology 144:3387-3395(1998).	
DR	EMBL: AJ005645; CAA06650.1;	
DR	InterPro: IPR001899; Gram_pos_anchor.	
DR	Pfam: PF00746; Gram_pos_anchor.1.	
DR	TIGRFAMS: TIGR01167; LPTXG_anchor.1.	
DR	TIGRFAMS: TIGR01168; YSIRK_signal.1.	
DR	PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.	
SO	SEQUENCE 947 AA; 102888 MW; 366FDF6E35121554 CRC64;	
Query Match	15.1%; Score 455; DB 2; Length 947;	
Best Local Similarity	27.2%; Pred. No. 6, 8e-10;	
Matches 158; Conservative 93; Mismatches 222; Indels 108; Gaps 25		
QY	35 GIEKRESDTESTYNDENATFLQTPDNNHLEEEVKESSSVES-----SNSIDTAQ	90
Db	45 GHEKRAEHTNGELNOSKEET-----APSEN-----KTKKVSRLQKDNVTCATND	92

```

QY 91 QPSHTTINRESVOTSDVNEHSHVDFANSKIKESNTESKEEINTIEOPNKVKEDSTISQ 150
DB 93 QPKVT-----MDSATVETSSNMOS-----PONATANOSTQ 125
QY 151 PSQVTINDEK---ISNODELLNLPINEYENKARPLSTTSAOPSIK-----VTVNOLAA- 201
DB 126 TSNTTNDKSTTYSNEDKSNL-----IOAKDVSTTPKTTTTPKPLNMAVNTAAAP 179
QY 202 EGGSNVHLIKVT--DOSITEGY--DSEGVIK--AHDAENILYDVEVDVKVSGDTMT 256
DB 180 QCGTNVDKHFNSNIDIAIDKGVNOTGKTEFWATSDVLKANTYIDDSVKEGDTFT 239
QY 257 VDIKNYPSDLTDSFTIPKIKDNGSELIATGYDNKKOQTYTFTPDVVDKEMKALK 316
DB 240 FKYOQFRRPGSVRLPSOTQNLNMOGNTIAKGIYDSTNTTFTTFTNVDYTVRGSFE 299
QY 317 LTSYIDSKVYNNNTKLDVEYKTALSSVNTIYEGORPNEENTANLOSMFTNIDTKNHT 376
DB 300 QVAFKRNKNTATTKTAYKMEVTLGNDTYSEIYDY--GNKKAQPLISSNTYINNEELS 356
QY 377 VEOITLYN-PLRYAKETNV-NISGNDEGSTIIDSTIIVKYGNQNL-----PDSN 429
DB 357 RNMTAYVNOQPKNTYTKQTFVTNLT-----GYKFNPNNAKFKIYEYTDQNGEVDSEFTPTDS 411
QY 430 RIYDYSEVEDYNDVYAOQLGNNNDVNFNGIDS--PYIIKVI-----SKYDPKMDYTT 482
DB 412 KLDVTDQFDVY---YSNDKNTATVLDLKKGOTSSKQIIOQVAVPDNSSTDNKIDY-- 466
QY 483 IOQVTVMQTTINEYTGEEFTASYDNTIAFTSSGOGGDLPEPKTYKIGDYVWEDVDKDG 542
DB 467 -----TLDTDKTKYSW---SNSYSVNGSSSTAND-----QKYNLGDYVWEDTJNKDG 511
QY 543 IQNTNNEKPLSNLVLTITYPDGTG--KSVRTDEGKIOFDG 562
DB 512 KODA--NEKGIGYVYILKDSNGKELDRTTDENGKYOFTG 550

RESULT 13
QY 053653 PRELIMINARY; PRT; 933 AA.
AC 053653;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
CX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEWMAN;
RC MEDLINE=94224142; PubMed=8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrinogen receptor
RT of Staphylococcus aureus."
RL MOL. MICROBIOL. 11:237-248(1994).
EMBL Z18852; CAJ79304.1;
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS00402; BPD_TRANS_P_NN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 933 AA; 97058 MW; EB51A6DEZFF759F4 CRC64;

Query March 15.08; Score 452; DB 2; Length 933;
Best Local Similarity 26.68; Pred. No. 8.7e-10;
Matches 154; Conservative 105; Mismatches 224; Indels 86; Gaps 26;
QY 1 SSDEKNDVYNNOSINTDNNQIIRKKEETNNYDGIERSDRETTSTNVDENATFLQK 60
DB 51 SNESKNSDSSSVSAAPKDDTNN-----SDTKSSNN--NGETSVAQ 91

```

```

QY 61 TPQONTHLTEEVEKSSSVSSNSIDTAQOPSHTTINRESVOTSDVNEHSHVDFANS 120
DB 92 NP-----AQETQOSSSTNATTEETPVGTGATTTTNOANTPATQOS--SNTMAEELVN- 143
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTTOSPSGVTINDEKISNODELLN-PINEYEKA 179
DB 144 --QTSNETTFDNTV-----SSVNSPONSTNAENVSTTQDTSTEATPSN---NES 189
QY 180 RPLSTTSAOPSIKRYTVN-----OLAAEGSNVHLIKVTDQ--SITEGYDSEGV 228
DB 190 APOSTDASNKDVNOAVNTSAPRRRAFELAAVADAPAGATDITNQLTNVTGVI--DSGTT 248
QY 229 IKADAENILYDVEVDVKVSGDTMTVDIDKNTVPESDLTDSFTIPKIKXNSGELITG 288
DB 249 VYRQAGVYKNTYGFSPNSAVKGDTEKITYPKELNLNGVSTAKVPPIMAGD--QVLANG 307
QY 289 TYDNKKQITTYTFTDYDYKYNIRKALKLTSYIDSKVYNNNTKL-DVEYKTALSS--VN 345
DB 308 VIDS-DGVNIYITFTDYVNTKDVAKATLTPAYID-----PEAVKKTGNTLATGIGSTTAN 362
QY 346 KTIYEGORPNEENTANLOSMFTNIDTKNHTVEQTIYNPL--RYSAKETNVNISGNDE 403
DB 363 KTVLVDYKGYKFFNLSTIKGTIIDIDKTNNTYRQTIYVNPESGDVYIAVLTLGNLKPNTDS 422
QY 404 GSTIIDSTIIVKYGNQNLPPDSNRIDYSEVEDYNDVYAOQLGNNNDVNFNG---- 459
DB 423 NALIDQQTSTIKYKVDNAADLSESYFV--NPENEDVNSVNTFTPNQYKVEENTPDD 481
QY 460 NIDSPYIIKVIYSKDPN-KDOYTTIQQVTVMQTTINEYTG--FRTASYDNTIAFTSSG 516
DB 482 QITPIYIVVNGHIDPNKGD-----LALRSTLYGNSNINIMSMWDEVAENNSG 534
QY 517 QGQG-DLP--PEKTYKIGDY--VWEDVDK-D-GIQNTNDN 549
DB 535 SGDGIDKPVVPEOPDEPEIPEPDSDDPSDGS DGS 573

RESULT 14
QY 0932C5 PRELIMINARY; PRT; 935 AA.
AC 0932C5;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Fibrinogen-binding protein.
GN FNB OR SAV0811.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
CX NCBI_Taxid=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Ihan J.-Q., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,
RA Katsura M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
EMBL AP003360; BAB56973.1;
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00402; BPD_TRANS_P_NN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;

```

Query Match 15.0% Score 452; DB 16; Length 935;

Best Local Similarity 25.9%; Pred. No. 8.7e-10;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```

QY 1 SSDEKNDVNNOSINTDDNNQIIKKEETNNYDIEKRSRDRFESTINVDENATFLQK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 51 SNEKSDNDSSVSAPKTDITN-----SPTKISSNTN---NETSYAQ 91

QY 61 TPQDNTLHEEVEKSSSVSSSIDTAQOPSHTTINRESVOTSDNVEDSHVDFANS 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 NP-----AQOETQSSSTNATTEETPVYTGATTTTQANTPATQS-SNTNAEELVN- 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTTSQPSGYTINDEKISNODELIN-LPINEYENKA 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 144 -QTSNETTSNDTNTV-----SSVNSPQNSTMAENVSTQDTSTEARPPSN---NES 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 RPLSTTSA-----OPSIKRVTVNOIAA-----EQSNVNHLIKVDQSITEGYDD 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 APQNTDASNKDVSOAVNPSTPRKAFSLAAVADAPAGTDITN--QITDVKVT---ID 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 SEGVIKADAEENLIYDTFEVDKVKSGDTMTVIDDKNTVPSDLTDSFTPIKIDNSGEI 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 SGTIVYPRQAGYVKNLNGFSVPNSAVGDTFKITVPKELNNGYSTAKVPPIMAGD-QY 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 IATGYDNKNQIYTFEVDYDKYENIKAKLTSYIDSKVPPNNNTKL-DVEYKTAAL-- 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 304 LANGVIDS-DGNVIYTFEVDYDNKENVTANTMPAYID---PEVWTKGTAVTLTGIGT 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 342 SSVKKTIVYQRPENENTANLOSMTNIDTKNHTVEGTIINPLRIYSAKETNV----- 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 359 NTASKTVLIDYEKGQFNLISIKGTIDQIDKTNNTYRQTIYVNP-----SGDNVLPALY 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 -NISNGDEGSTIIDSTIIKVKYKGNQNL.PDSNRIDYSEYEDVTNDYDQAOLGNNDV 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 414 GNLPNTKSNALIDAKNTDIVRYR-DNANDLSESYVNPDEFEDVTNQVRISEPNANQY 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 NINF-----GNIDSPYIIKIVISKYDPNKDDYTTIQOVTMTQTTINEYTGEF--RTASYDNT 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 473 KVEPPTDDQITTPYIVVNGHIDPASTG-----DLALRSTFYCYDSNFIIRSMWNE 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 509 IAFSTSSGQGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 549
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 VAFNNGSGSGDIDKPVVPEQDEGEIEPIPEDSDSPGSDSGSDS 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

Q99VJ4 PRELIMINARY; PRT; 989 AA.

```

AC Q99VJ4;
T 01-JUN-2001 (Tremblrel. 17, Created)
01-JUN-2001 (Tremblrel. 17, Last sequence update)
01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fibrinogen-binding protein A, clumping factor.
GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsunaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-U Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kihara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003131; BAB41975.1;

```

DR InterPro: IPR000515; BPD_transp.

DR InterPro: IPR001899; Gram_pos_anchor.

DR TrGFams; TrGR01167; LpxTG_anchor; 1.

DR PROSITE; PS00402; BPD_TRANS_PNN_MBR; UNKNOWN_1.

DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.

KW Complete proteome.

SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 15.0% Score 452; DB 16; Length 989;

Best Local Similarity 25.9%; Pred. No. 9.2e-10;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```

QY 1 SSDEKNDVNNOSINTDDNNQIIKKEETNNYDIEKRSRDRFESTINVDENATFLQK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 51 SNEKSDNDSSVSAPKTDITN-----SPTKISSNTN---NETSYAQ 91

QY 61 TPQDNTLHEEVEKSSSVSSSIDTAQOPSHTTINRESVOTSDNVEDSHVDFANS 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 NP-----AQOETQSSSTNATTEETPVYTGATTTTQANTPATQS-SNTNAEELVN- 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTTSQPSGYTINDEKISNODELIN-LPINEYENKA 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 144 -QTSNETTSNDTNTV-----SSVNSPQNSTMAENVSTQDTSTEARPPSN---NES 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 RPLSTTSA-----OPSIKRVTVNOIAA-----EQSNVNHLIKVDQSITEGYDD 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 APQNTDASNKDVSOAVNPSTPRKAFSLAAVADAPAGTDITN--QITDVKVT---ID 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 SEGVIKADAEENLIYDTFEVDKVKSGDTMTVIDDKNTVPSDLTDSFTPIKIDNSGEI 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 SGTIVYPRQAGYVKNLNGFSVPNSAVGDTFKITVPKELNNGYSTAKVPPIMAGD-QY 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 IATGYDNKNQIYTFEVDYDKYENIKAKLTSYIDSKVPPNNNTKL-DVEYKTAAL-- 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 304 LANGVIDS-DGNVIYTFEVDYDNKENVTANTMPAYID---PEVWTKGTAVTLTGIGT 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 342 SSVKKTIVYQRPENENTANLOSMTNIDTKNHTVEGTIINPLRIYSAKETNV----- 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 359 NTASKTVLIDYEKGQFNLISIKGTIDQIDKTNNTYRQTIYVNP-----SGDNVLPALY 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 -NISNGDEGSTIIDSTIIKVKYKGNQNL.PDSNRIDYSEYEDVTNDYDQAOLGNNDV 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 414 GNLPNTKSNALIDAKNTDIVRYR-DNANDLSESYVNPDEFEDVTNQVRISEPNANQY 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 NINF-----GNIDSPYIIKIVISKYDPNKDDYTTIQOVTMTQTTINEYTGEF--RTASYDNT 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 473 KVEPPTDDQITTPYIVVNGHIDPASTG-----DLALRSTFYCYDSNFIIRSMWNE 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 509 IAFSTSSGQGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 549
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 VAFNNGSGSGDIDKPVVPEQDEGEIEPIPEDSDSPGSDSGSDS 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: March 14, 2003, 13:00:05
Job time : 27.2509 secs